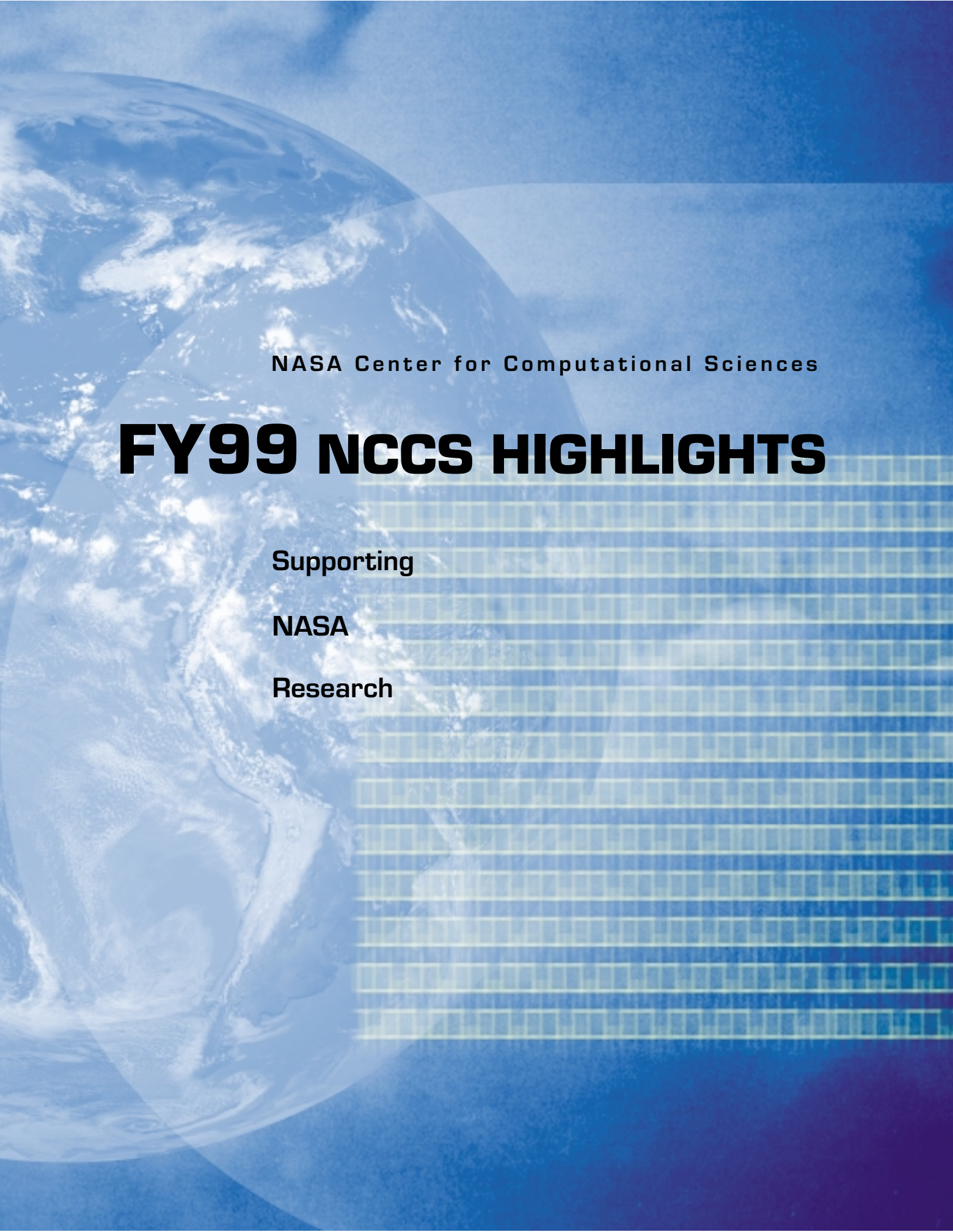


NASA Center for Computational Sciences

FY99 NCCS HIGHLIGHTS

Supporting

NASA

Research

FY99 NCCS HIGHLIGHTS

Supporting NASA Research

PREFACE

The NASA Center for Computational Sciences (NCCS) is a high-performance scientific computing facility operated, maintained, and managed by the Earth and Space Data Computing Division (ESDCD) of NASA Goddard Space Flight Center's (GSFC) Earth Sciences Directorate. The mission of the NCCS is to advance leading-edge science by providing the best people, computers, and data storage systems to NASA's Earth and space sciences programs and those of other U.S. Government agencies, universities, and private institutions.

Among the many computationally demanding Earth science research efforts supported by the NCCS in Fiscal Year 1999 (FY99) are the NASA Seasonal-to-Interannual Prediction Project, the NASA Search and Rescue Mission, Earth gravitational model development efforts, the National Weather Service's North American Observing System program, Data Assimilation Office studies, a NASA-sponsored project at the Center for Ocean–Land–Atmosphere Studies, a NASA-sponsored microgravity project conducted by researchers at the City University of New York and the University of Pennsylvania, the completion of a satellite-derived global climate data set, simulations of a new geodynamo model, and studies of Earth's torque.

This document presents highlights of these research efforts and an overview of the NCCS, its facilities, and its people.

CONTENTS

1 **Enabling Earth Science**
The Facilities and People of the NCCS
The NCCS is a high-performance scientific computing facility operated, maintained, and managed by the Earth and Space Data Computing Division of GSFC's Earth Sciences Directorate.

12 **NASA's Seasonal-to-Interannual Prediction Project**
In Partnership With the NCCS
Researchers with NASA's Seasonal-to-Interannual Prediction Project study El Niño and related phenomena using the NCCS's supercomputers to run their computationally intensive models.

20 **Low-Level Jets**
The Data Assimilation Office and Reanalysis
Through its partnership with NASA's Data Assimilation Office, the NCCS supports scientists who use the technique of data assimilation to gain a better understanding of Earth processes.

28 **North American Observing Systems**
An Interagency Group Runs Tests at the NCCS
National Oceanographic and Atmospheric Administration (NOAA) scientists use NCCS resources to test the effectiveness of weather observing systems by simulating the removal of some of the launching sites in the Nation's weather balloon network, testing redundancy with measurements taken by aircraft.

36 **Climate Prediction Sees Future Despite Chaos**
Researchers Outside NASA Use NCCS Resources for Studies
Climate predictions attempt to forecast the weather in broad strokes, months or more in advance. Climatologists from an outside organization use NCCS computers with the aim of predicting climate changes and achieving greater than 50 percent accuracy.

42 **TOVS Pathfinder Path A**
A Boon for Climatologists
The TOVS Pathfinder Path A, a 21-year data set based on measurements taken from several NOAA satellites, was created by a researcher at GSFC using the facilities at the NCCS.

48 **An Interview With Dr. Joanne Simpson**
Her Career, Meteorology, and Computers
Dr. Simpson is sometimes called the Marie Curie of meteorology. Her long and distinguished career includes the past several years as Chief Scientist for Meteorology at GSFC.

58 **Microgravity**
Molecular Dynamics Simulations at the NCCS Probe the Behavior of Liquids in Low Gravity
A speck of water falls on a piece of Teflon, and surface tension causes it to bead. Two university researchers are using NCCS computers to probe this type of behavior.

64 **The Torque of the Planet**
NASA Researcher Uses NCCS Computers To Probe Atmosphere–Land–Ocean Coupling
A NASA researcher has been modeling the effect of atmospheric torques of various origins using the NCCS's Cray J932se.

70 **Synthetic Aperture Radar**
The NCCS Enables Search and Rescue
NASA's Search and Rescue Mission has been developing a new tool to aid in rescuing victims of aviation disasters, an effort supported by the NCCS.

78 **The Earth Gravitational Model 1996**
The NCCS—Resource for Development, Resource for the Future
A mathematical model of Earth's gravitational field allows oceanographers to extract ocean current information from satellite altimeter data.

84 **Simulating the Dynamics of Earth's Core**
Using NCCS Supercomputers Speeds Calculations
Only in the last few years has it been possible to study accurately Earth's core, by developing a mathematical model and performing simulations on supercomputers such as those at the NCCS.

90 **Appendixes**
FY99 NCCS-Supported Research Projects and Principal Investigators
Acronyms, Names, and Abbreviations
Researchers' Affiliations

ENABLING EARTH SCIENCE

The Facilities and People of the NCCS

Photo of flags by Jason Bardi. GOES image of Earth courtesy of NASA.

The NCCS provides computing resources for nearly 200 research efforts.

The mission of the NASA Center for Computational Sciences (NCCS) is to advance leading-edge science by providing the best people, computers, and data storage systems to NASA's Earth and space sciences programs and those of other U.S. Government agencies, universities, and private institutions.

Technologies for acquiring, storing, processing, analyzing, managing, distributing, and visualizing massive amounts of data are evolving rapidly, as are the needs of the scientists who use them. To meet these needs, the NCCS's supercomputers, high-speed networks, and mass storage facilities also must evolve, keeping at the forefront of technology. Co-evolving with the facilities are the people of the NCCS, who support the NCCS's computer facilities and provide their expertise to Earth and space sciences researchers.

The NCCS supports primarily Earth science efforts, as well as flight systems projects, space science research, and Hubble Space Telescope (HST) activities.

More than 700 U.S. scientists have accounts on the NCCS's supercomputers. These state-of-the-art systems allow scientists to perform sophisticated, large-scale modeling of Earth's gravity field, land–ocean–atmosphere feedback, 3–D cumulus cloud models, and numerical climate prediction, for example.

The NCCS's mass data storage system allows scientists to store and manage the vast amounts of data generated by these computations, and its high-speed network connections allow the data to be accessed quickly from the NCCS archives. Some NCCS users perform studies that are directly related to their ability to run computationally expensive and data-intensive simulations. Because the number and type of questions scientists research often are limited by computing power, the NCCS continually pursues the latest technologies in computing, mass storage, and networking technologies.

Just as important as the processors, tapes, and routers of the NCCS are the personnel who administer this hardware, create and manage accounts, maintain security, and assist the scientists, often working one on one with them.

History of the NCCS

Less than a year after President Dwight D. Eisenhower established the National Aeronautics and Space Administration (NASA) on July 27, 1958, NASA created the Goddard Space Flight Center (GSFC) in Greenbelt, Maryland, to be the primary receiving and processing center for satellite data. As an increasing number of satellites began imaging Earth's atmospheres and oceans, GSFC became a center for Earth science researchers.

To support these researchers, GSFC's Earth Sciences Directorate purchased a Control Data Corporation CYBER 205 supercomputer, with 32 megabytes (MB) of central memory and 10 gigabytes (GB) of disk storage, and established the NCCS as part of the Earth Science Computing Division. An IBM 3851 storage system and two Masstor M860 magnetic tape cartridge units provided an additional 375 GB of mass data storage.

Since 1984, the NCCS has continually upgraded its supercomputing and mass data storage systems to accommodate the increasing computing requirements of NASA's science community. Some major technological changes have occurred over that time, such as the advent of massively parallel computing, the birth of the Internet, the growth of powerful desktop computing, and a move away from centralized computing for all but the largest of applications.

As workstations and powerful desktop systems came into use, the NCCS's user community decreased from about 1,500 heavily used accounts to about 700 accounts today. However, the need for larger simulations and increased speed continued to grow, as did the role of the NCCS. The NCCS now focuses primarily on high-end computing, with several hundred heavily used accounts.

The NCCS is a centralized high-end computing facility staffed by scientists, programmers, and information technology professionals.

At the end of FY99, the NCCS's two Cray J932se's, an SGI Origin 2000, and a Cray T3E had a combined total of 1,152 processors, 178 GB of central memory, and 3,629 GB of disk storage. The NCCS's Mass Data Storage and Delivery System (MDSDS) had more than 1 petabyte (PB) of potential near-line data storage—the equivalent of 1 billion floppy disks.

The NCCS was formed in 1984 to support Earth and space science research.

Facilities

COMPUTING

The NCCS provides two types of architecture: vector and parallel. Committed to maintaining and supplying some of the fastest vector machines possible for its users, the NCCS maintains a pair of Cray J932se supercomputers. Each achieves speeds of 6.4 gigaFLOPS (GFLOPS) and have 8 GB of main memory. One has 1,260 GB of disk storage and the other has 450 GB.

Both machines are connected to two StorageTek (STK) Automated Cartridge System silos with a Powderhorn robot and six STK Timberline 9490 cartridge tape drives, and to a Wolfcreek robot with two STK Timberline 9490 cartridge tape drives. The total storage capacity for the cluster is 7.6 terabytes (TB) compressed (or 5.0 TB uncompressed). Both operate on the SGI UNICOS 10.0.0.1 operating system, with C, C++, and Fortran 90 compilers. Available libraries are BLAS, EISPACK, FISHPACK, HDF, IMSL, NAG, ODEPACK, and SLATEC; available utilities are FLINT, prof, perftrace, hpm, atexpert, and totalview.

The NCCS has a parallel processing SGI Origin 2000, which supports research efforts at the Data Assimilation Office (DAO). It has 64 R10000 processors with cache-coherent Non Uniform Memory Access (ccNUMA) architecture, 8 GB of main memory, and 449 GB of disk storage. It operates on the SGI Irix 6.5 operating system and has C, C++, FORTRAN 77, and Fortran 90 compilers installed.

Cray T3E

Finally, the NCCS and ESDCD's High Performance Computing and Communications (HPCC) program share one of the world's fastest parallel supercomputers. The Cray T3E, is the workhorse of the NASA Seasonal-to-Interannual Prediction Project (NSIPP). With 1,024 processors, it is able to achieve 613 GFLOPS; it has 130 GB of memory and 1,470 GB of disk storage. Attached is an STK Automated Cartridge System silo running 8 STK Timberline 9490 cartridge tape drives with 600 800-MB-capacity tapes, and 4 STK Redwood cartridge tape drives with 750 10-GB tapes, 680 25-GB tapes, and 350 50-GB-capacity tapes, for a total capacity of 63.7 TB compressed (42.5 TB uncompressed).

MASS STORAGE

Mass storage is the natural complement to high-power computing because the output from some computer model simulations can be as large as a terabyte. Output of that size is routinely broken down into more manageable units for ease of retrieval—creating even more files. Accommodating today's growing data storage needs is a top priority of the NCCS. Never before has the science world produced such massive

quantities of data. Today there are more than 6 million files containing more than 100 TB of user data under the control of the NCCS MDSDS.

The MDSDS operates under a Sun E10000 server, which uses UniTree Central File Manager Version 2.1, from UniTree Software, Inc. (UTSI). UniTree is an intelligent hierarchical data archival system initially developed at Lawrence Livermore Laboratories in the 1980s that has the ability to run on several different UNIX platforms. Files written to the server are copied automatically onto tapes while still remaining on disk. Less frequently used data files are purged from disk when space is needed and transparently copied back from tape to disk when the file is read again. This allows the most frequently used data to remain on disk for faster access.

The NCCS's UniTree system is one of the largest and most active systems in the unclassified world.

With a disk cache of 1.5 TB, this machine is connected to 56 tape drives in 7 silos and an IBM 3494 robotic library, as well as 4 freestanding Timberline drives. The computer has 12 CPUs, 8 GB of memory, 12 GB of system disk storage, 817.5 GB of EMC disk storage, 775 GB of CLARiiON disk storage, and 196 TB of robotically managed storage.

Another milestone was reached on September 7, 1999, when the total amount of data stored on the MDSDS exceeded 100 TB, setting a record among UniTree sites in the unclassified world.

The MDSDS set three new records in July 1999: The largest amount of total data transferred to UniTree in a day—302.927 GB on July 16. The highest total stored in a week—1,152 GB the week of July 25. A record for new data added to UniTree in a month—4.1 TB.

TAPE LIBRARIES AND DATA PROTECTION

Seven STK 9310 Powderhorn robotic storage silos constitute the bulk of the near-line storage of data under MDSDS. One of these silos uses 6 STK Redwood tape drives, and the remaining 6 silos use 8 Timberline tape drives and 30 STK 9840 Eagle tape drives.

An IBM 3494 Tape Library Dataserver with 16 IBM 3590 Magstar tape drives uses 4,200 10-GB Magstar tapes, for a total uncompressed capacity of 41 TB. Operator-mounted storage includes 4 STK Timberline tape drives and 20,000 cartridge tapes in the freestanding, manually mounted tape library, with a capacity for 128,800 3490E cartridge tapes—approximately 128 TB.

To reduce the risk of data loss from tape cartridge failure, the NCCS provides duplicate data storage at a remote location. In 1997, the NCCS installed a duplicating complex to write and store copies of all files controlled by the MDSDS. This complex

The total potential storage capacity for the NCCS mass storage system is over 1 PB.

consists of a StorageTek silo with six StorageTek Redwood drives that can stream large files at up to 12 MB/sec. Three thousand 25-GB and 500 50-GB Redwood cartridge tapes reside in the remote silo. All data files created since the installation of the remote silo have been duplicated automatically, and NCCS personnel have been duplicating the 40 TB of pre-existing data.

NETWORKING

NCCS users have transferred as much as 300 GB to the MDSDS in a single day. With this volume of network traffic, the NCCS must maintain a full complement of networking capabilities to support the wide-ranging networking needs of all users, whether they work at GSFC or access the NCCS facilities remotely.

Users on the GSFC campus have a range of networking options available to them. The NCCS supports Transmission Control Protocol/Internet Protocol (TCP/IP) traffic over Ethernet, with speeds up to 10 Mb/sec; Fast Ethernet TCP/IP, with speeds up to 100 Mb/sec; Fiber Distributed Data Interface (FDDI), with speeds up to 100 Mb/sec; Asynchronous Transfer Mode (ATM) OC3, with speeds up to 155 Mb/sec; and ATM OC–12, with speeds up to 622 Mb/sec. In early 2001, campus access to the NCCS from workstations having gigabit Ethernet network connections will be provided via the new Science ATM Network (SAN).

Users outside GSFC also have several options for accessing NCCS systems. Generally, most remote users are routed through a wide-area network T3 line, which has a standard transfer rate with speeds up to 45 Mb/sec—nearly 50 times faster than some cable modems. From some NASA centers and research partner organizations, the NCCS can be accessed at speeds up to 100 Mb/sec by means of the NASA Research and Engineering Network (NREN). If outside users need a faster connection, they can be routed through the University of Maryland Mid-Atlantic Crossroads (MAX), which maintains a wide-area ATM OC–12 link to the ABILENE Internet 2 research network and has a direct line to NASA GSFC at the 622 Mb/sec rate.

Finally, the internal connections between the supercomputers and the mass storage system are some of the fastest currently possible. The Crays and the MDSDS communicate through two High-Performance Parallel Interface (HiPPI) switches, which have a data transfer rate of up to 800 Mb/sec—about 15,000 times faster than a standard 56k modem.

UPGRADES IN FY99

Cray J932se's

The NCCS traded in one of its three Cray J932se supercomputers for an SGI/Cray Origin 2000 system, with 64 CPUs, 32 GB of memory, and 1,587 GB of disk storage. This system will be used exclusively for two priority DAO projects: Retrospective Analysis Research and Statistical Digital Filter Research. The Origin 2000 will provide the NCCS with experience in optimizing code for ccNUMA architecture and will enable the DAO to develop retrospective analysis codes.

In the future, the other two Cray J932se's will be upgraded to SV1s, exchanging 32 existing se processors for 24 SV1 processors. In FY99, this upgrade was approved as a way of supporting the general user community. According to SV1 benchmarks, Earth science codes ranged from 230 percent to 310 percent faster than on the current J90se processor.

In 1999, the NCCS received and installed an IBM SP testbed platform for benchmarking code created by the DAO, NSIPP, and other GSFC Earth science programs to help the NCCS make future purchasing decisions.

Mass Storage System

The mass storage system at the NCCS underwent many changes in FY99. The NCCS upgraded tape drives, silo control software, and UniTree software, and twice upgraded the mass storage platform itself.

Mass Storage Server and Disk

After 5 years of running the MDSDS using HP–Convex/UniTree on an HP C3830 with 3 CPUs, 2 GB of memory, and up to 555 GB disk cache, the NCCS began running (on January 27, 1999) UniTree software on a Sun E6500 with 8 CPUs, 8 GB of memory, and more than 800 GB disk cache. The new system had more than three times the input/output (I/O) capacity of the older system at one-twelfth of the maintenance cost.

With user storage needs expected to increase considerably, a further upgrade was executed to give the MDSDS more disks. On June 23, 1999, the NCCS transferred control of the MDSDS hardware platform from the Sun E6500 to a 12-processor Sun E10000. With nearly five times the I/O bandwidth, 4 additional CPUs, 11 additional SCSI adapter slots, and room for 4 additional system boards, the E10000 gives the NCCS the ability to scale up to whatever the near term might require. The E10000 has 36 empty CPU slots available for expansion, 477 percent more data bandwidth than the E6500 (12.8 GB/sec vs. 2.68 GB/sec bandwidth), and room for 4 additional system boards, each of which can have 4 additional SCSI slots and 4 additional CPUs. To take advantage of the E10000, the NCCS purchased 1.3 TB of EMC disk space and 900 GB of StorageTek CLARiiON disk space.

Most important, the E10000 can run multiple system images on the same machine, which will allow the testing of operating system patches and upgrades without putting the production environment at risk.

The NCCS has retained the E6500 to support both a data warehousing research project and an NCCS-written utility that allows data to be read from and written to tape. Other uses of this system are anticipated in the coming year.

Tape Drives

The NCCS replaced its tape drive technology with newer, faster, and denser drives and tapes. Three years ago, MDSDS had 40 block-mux tape drives, capable of reading and writing a maximum of 3 MB/sec and filling a tape with up to 800 MB of uncompressed data. At the start of FY99, MDSDS used 24 StorageTek Timberline drives, capable of reading and writing up to 6 MB/sec, as well as 16 IBM Magstar tape drives, capable of speeds up to 9 MB/sec and filling a tape with 10 GB of uncompressed data. These IBM drives were in a separate 3494 robotic library. In February and March, the NCCS installed 30 StorageTek 9840 drives, capable of performing I/O at 9 MB/sec and filling a tape with up to 20 GB of uncompressed data, and reduced the number of Timberline drives to 12.

User Support Services

One of the best and most important services that the NCCS offers is seamless support. The NCCS is staffed with experienced professionals who maintain its resources and support its scientific user community in everything from logging on to running code to saving and retrieving files. The NCCS is committed to staffing its support groups with highly qualified personnel with backgrounds in information technology and scientific computing.

The NCCS User Administrator is often the first point of contact for outside users, providing assistance to NCCS division representatives, sponsors, and new users with questions, problems, comments, and suggestions. The user administrator also processes requests from NASA Headquarters funding managers and from GSFC Directorate or Division funding personnel. Funding for a scientific research effort comes from NASA Headquarters or through a GSFC Directorate or Division, which authorizes the computing time.

TECHNICAL ASSISTANCE GROUP

The Technical Assistance Group (TAG) staff, some of whom have degrees in meteorology, physics, mathematics, or geography, help the researchers use the computers and mass storage systems more effectively.

Users send questions and problem reports to the help desk through a variety of mechanisms such as E-mail, telephone, phonemail, and the World Wide Web (WWW). In FY99, the TAG received nearly 800 requests for help-desk support. Typical questions concern debugging and optimizing Fortran and C programs, using systems and applications software, solving I/O and I/O media-related problems, backing up and restoring files, and using other installed utilities and NCCS facilities. While the TAG team resolves most user issues, it also can refer users to the Applications Support Group for one-on-one consulting (see page 8).

The NCCS tailors its training services to support users during transitions to new hardware and software systems and to help them use NCCS resources as productively as possible. NCCS personnel regularly teach short classes on a number of subjects of interest to the scientists who use these facilities, including Fortran 90 and basic opti-

mization, covering data types, procedures/modules, I/O, conversion issues, and vectorization/parallelization.

Documentation is available by user request or through Internet access. The NCCS WWW documentation server provides general information about the NCCS and its resources, technical information about NCCS systems, system status information, news articles, and minutes of Computer Users Committee (CUC) meetings.

APPLICATIONS SUPPORT

One of the biggest challenges in large-scale computing is optimizing the software. In addition to new code, many algorithms in use today were designed decades ago on older machines, and have evolved over time, sometimes in an ad hoc fashion. The NCCS offers one-on-one consulting through the Applications Support Group to help users optimize codes to run on NCCS's vector machines and the Cray J932se's. Staff members also help convert vector codes to parallel codes, thus improving the algorithm's execution so that it makes the best use of the NCCS's Cray T3E. The goal is to allow scientists to complete their research more rapidly and to increase the size of problems they can tackle.

- The single-CPU optimization of a thermal dynamic ocean model for GSFC's Antonio Busalacchi, of Code 970's Laboratory for Hydrospheric Processes, was a challenge. The effort was a success—the code's performance on the Cray J90 improved from 38 MFLOPS to 58 MFLOPS, and the 2-year simulation time was reduced from 30 to 15 hours. This work continues into FY00 with the ultimate goal of running the model five times faster by multitasking it.
- The scaling of the Climate and Radiation Branch's 2–D Goddard Cumulus Ensemble Model was optimized in FY99. Previously, the model was run on only one CPU and took up to 18 hours of wall-clock time to perform a 24-hour simulation. The same simulation can now be done in less than 2 hours of wall-clock time. This optimization has enabled Chung-Hsiung Sui, of GSFC's Laboratory for Atmospheres, Climate and Radiation Branch, to expand his research by including more cases and using a larger model domain.
- Holding discussions with the Canadian Meteorological Center (CMC), the European Center for Medium Range Weather Forecasts (ECMWF), and the National Centers for Environmental Prediction (NCEP), the Applications Support Group completed a feasibility study for the DAO. Its purpose was to determine if systems from these centers could be used as backup for the Earth Observing System (EOS) AM–1 launch. The study concluded that the NCEP system could be used and would produce the least amount of risk.

The NCCS offers one-on-one consulting through the Applications Support Group.

One of the major advantages of having a centralized computing facility like the NCCS is that it enables efficient maintenance of systems and security.

- The Applications Support Group began reengineering code for Wei-Kuo Tao of GSFC's Mesoscale Atmospheric Processes Branch to enable it to run on parallel platforms. Tao's code was originally written in FORTRAN 77 to run on vector machines. Using an inhouse tool for domain decomposition, staff began to adapt the code for a parallel environment, enabling interprocessor communication and rebuilding the code from the ground up, routine by routine.

SYSTEMS SUPPORT

Systems staff maintains the machines, keeping them on line and installing system upgrades when necessary, and serving as administrators of the various machines.

A major function of the systems administrators is to maintain security. The focus of the NCCS security team is on the organizations and projects within the domain of the computing center and related projects and facilities. The security team is also part of a grass-roots, centerwide security organization that benefits users and administrators in the NCCS and elsewhere—the NCCS makes available its security experts to consult for other organizations within GSFC. The security team's chief challenge is finding a way to balance the need to build and maintain effective security with the need to provide public services to the users. An open GSFC network environment is essential to the productivity of the scientific community; however, a significant effort is required to protect NCCS computing resources in this type of environment.

COMPUTER USERS COMMITTEE (CUC)

The NCCS constantly strives to better match its environment to the changing needs of its users and provide the best services and facilities, such as upgrading its systems to stay current with evolving technology.

The NCCS formed the CUC whose members provide input to NCCS decisions about computer hardware and software. The committee also forms subcommittees to address specific issues, such as Cray utilization or the UniTree mass data storage system, meeting bimonthly to obtain input from customers, promote communication between the NCCS staff and the user community, announce changes, and work to improve the facilities and their operation.

In April 1999, the NCCS met with its user community to gain feedback on queue structure improvements, disk quota alterations, and other topics that would improve usage. One result of this meeting was a decision by the NCCS to accommodate larger jobs both in memory and in CPU time. For example, users with multitasked jobs may now request up to 480 Megawords (MW) of memory, whereas previously they were limited to 280 MW. These changes will allow users to solve problems that are larger and at finer resolutions.

The NCCS believes that, in order to provide a high level of support, its staff must be actively involved in the technology. The lead technical person for the NCCS's mass storage system is also president of the UniTree Users Group. This group was formed in 1992 to help determine whether the needs of UniTree users across the country are being met and to identify suggestions for system enhancements and developments.

The Future

One of the greatest challenges in the coming years will be to meet the growing data needs of the Earth science modeling community. Aside from maintaining some of the best machines, there are several definite goals. Supporting software design is one of these.

The NCCS is committed to bringing in highly talented people with backgrounds in computer science and software engineering to optimize code for the supercomputers and to help the scientists manage their codes' long-term evolution—work beginning in the year 2000. Modernization can be as simple as converting code from FORTRAN 77 to Fortran 90, but sometimes modernization implies more extensive object-oriented software design. The NCCS knows that scientists often have neither the time nor the resources to perform the modernization themselves. The NCCS seeks to work with them to reengineer their codes while maintaining performance and fostering overall better software development, allowing the scientists to focus on the science.

NASA'S SEASONAL-TO-INTERANNUAL PREDICTION PROJECT

In Partnership With the NCCS

Researchers with NASA's Seasonal-to-Interannual Prediction Project (NSIPP) refer to different types of memory when running models on NCCS computers: the computer memory required for their models and the memory of the atmosphere or the ocean.

Because of the atmosphere's chaotic nature, its memory is short. For weather predictions, the initial information taken from atmospheric observations has a limited useful life. Currently, there is no way to take observations, initialize an atmosphere model, integrate ahead in time, and make an accurate weather forecast beyond about 2 weeks. After that, the system becomes chaotic.

What conditions could be used to make predictions beyond 2 weeks? If not conditions in the atmosphere, then the memory must be found somewhere else. That place is in the oceans.

Although most changes in the atmosphere vary on a short timescale, the weather being a prime example, some important large atmospheric climate variations occur over much longer timescales—months, years, or decades. NSIPP is interested specifically in those phenomena that occur over timescales of several months to a few years, and the El Niño Southern Oscillation (ENSO) is the most significant of these.

ENSO encompasses both the periodic warming El Niño and its cooling counterpart, La Niña. In El Niño events, the large-scale condensation and heat release normally associated with the western Pacific shift eastward, disrupting normal atmospheric

View of the tropical Pacific, courtesy of NOAA.

For its studies, NSIPP uses a Cray T3E called jsimpson, which the NCCS owns jointly with NASA's High-Performance Computing and Communications Program (HPCC).

circulation over the ocean. Atmospheric winds force the ocean waters to move, and this motion shifts the location of warm waters, which affects the direction and strength of winds, which move the waters, and so on (see "El Niño and Nature's Feedback"). "The NSIPP program," says GSFC's Michele Rienecker, "is predicting the phase and amplitude of patterns in the eastern equatorial Pacific and their effects on the midlatitudes."

NSIPP has a close relationship with the NCCS, which has supported NSIPP work for a number of years (see "NSIPP–NCCS: A Partnership"). This close relationship is important as NSIPP relies more and more on massive computers in the development and testing of its models and as it moves toward making operational predictions of El Niño events and their impact.

The importance of El Niño events lies in their profound impact on weather around the globe. NSIPP scientists have turned to the oceans to understand and predict El Niño. "The tropical ocean," says NSIPP's Max Suarez, "is what's really critical for the El Niño prediction."

The circulation of the oceans is somewhat chaotic, like that of the atmosphere, but, because water has a much higher heat capacity than air, oceans cycle temperature changes more slowly than does the atmosphere. Nowhere is this property more important than on the tropical Pacific Ocean's thermocline—the subsurface interface separating warm surface waters from cooler, nutrient-rich waters in the ocean's depths. Inertia in this thermocline gives the ocean its memory.

It takes much longer for waves to propagate back and forth across a distance on the thermocline than across the same distance in the atmosphere. This is especially true in the wide Pacific Ocean, where waves on the thermocline can take as long as 3–9 months to cross—as opposed to the days or weeks it might take for a weather system to traverse the same distance.

NSIPP–NCCS: A PARTNERSHIP

"If you are doing a serious large-scale model, you need a state-of-the-art computer system," says Rienecker. Moving to the T3E was, according to Rienecker, a big boost for the NSIPP program because the computer offers increased scalability and has been essentially dedicated for NSIPP use. NSIPP personnel have been able to run more simulations and experiments.

In FY99, the T3E boasted 1,024 processors, of which half were dedicated to NSIPP work.

"Working with the NCCS to configure the T3E has really helped us," says Rienecker. The partnership started in 1997 when NSIPP petitioned NASA Headquarters and the ESDCD to purchase the T3E from the German Meteorological Institute and dedicate it to NSIPP. NSIPP and the NCCS worked together to determine the optimal usage of the machine, ascertaining the optimum size and retention period of disks, finding the best way to migrate to the hierarchical system, taking note of file sizes, setting up different groups within the UniTree system, and determining how best to access files. The NCCS further assisted in communicating all of this information to the scientists involved. "[The partnership with the NCCS] makes a tremendous difference in how we use the machine efficiently," says Rienecker.

Rienecker has been pleased with the results so far. "It's a pleasure to use these systems," she says of the Cray T3E and the NCCS mass storage system.

These longer timescales are what enable NSIPP to make long-term predictions.

One of NSIPP's goals is to determine the effects of "teleconnections" between ENSO and areas outside the Tropics: How does El Niño affect the weather in North America, for instance? How is El Niño information transmitted through the atmosphere? How will the warming of waters in the Pacific affect the jet stream?

El Niño events occur at irregular intervals as short as 2 years to as long as 7 years, and with varying duration. The 1997 El Niño lasted only a year, for example, while the early 1990's saw an El Niño that lasted almost 3 years. The economic consequences of an El Niño can be staggering. As a typical example, more than $33 billion in damages can be attributed directly to the 1997 El Niño alone. Predicting an event might prevent some of the damage (see "Why Predict

EL NIÑO AND NATURE'S FEEDBACK

El Niño is the name given to the occasional warming of surface waters in the central and eastern equatorial Pacific Ocean. Normally, the prevailing trade winds of the Pacific blow east to west along the surface at the Equator. These winds pile warm water in the upper ocean of the western tropical Pacific, near Indonesia and Australia, and this water heats the atmosphere, creating convection and precipitation. The convection drives the large-scale atmospheric "Walker" circulation, which is the name given to the combined surface east-to-west trade winds and west-to-east winds in the upper levels of the atmosphere.

The name El Niño, which means "The Boy" or "The Christ Child," was coined by Peruvian fishermen because the events would usually occur around Christmas.

At irregular intervals—about every 3–5 years, but sometimes as frequently as every 2 or as infrequently as every 7 years—an El Niño occurs. The normal trade winds relax, and the warm pool of water in the western Pacific shifts back along the Equator toward South America. The warmest water in the Pacific—typically about 30°C around Indonesia—then shifts eastward. This causes the normally cooler waters of the eastern Pacific, off the coast of South America, to warm. The warmest pool moves to the center of the tropical Pacific, where it causes condensation, heat release, and atmospheric upwelling. This upwelling disrupts the normal Walker circulation over a period of months.

These changes have a profound effect on the atmosphere. The warm pool in the central and eastern Pacific usually causes heavier than normal rains over tropical South America, while the mechanism for precipitation in the western Pacific is shut off. Ecuador and Peru may flood, while Indonesia, Australia, and Zaire may have droughts. The jet stream into North America becomes stronger, and an increased amount of moisture is carried into the southern States. The atmosphere also produces alternating patterns of low- and high-pressure systems. Typically, a low-pressure system centered just to the southwest of Alaska draws warm air up into Canada and creates a tendency for higher than normal temperatures in western Canada and the upper plains of the United States. A low-pressure system centered over the southeastern United States draws cold moist air into that region and brings lower than normal temperatures to the South.

This same low-pressure system also increases precipitation in areas around the Gulf of Mexico and slows the mechanism for forming hurricanes. There are usually fewer hurricanes during an El Niño year.

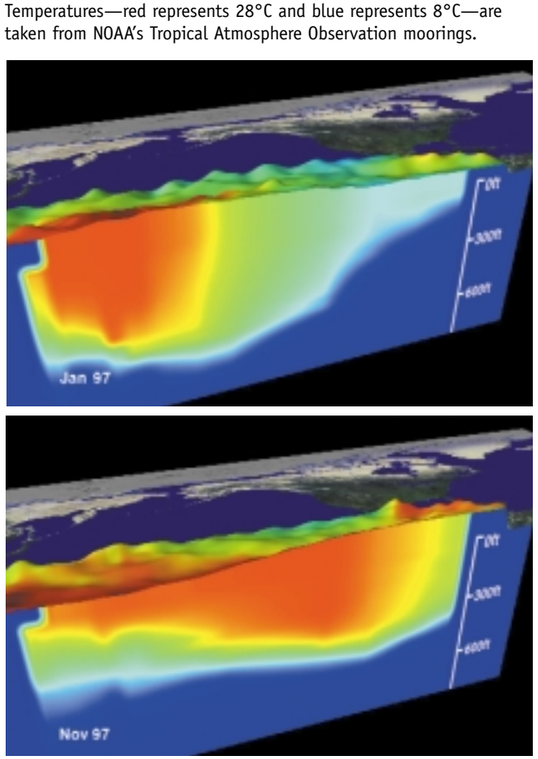
A portion of the 1997–1999 ENSO event with SST superimposed upon SSH, courtesy of Gregory Shirah, NASA.

El Niño?"), but, because of the chaotic nature of the atmosphere, exact (i.e., deterministic) teleconnections cannot be predicted.

However, scientists at NSIPP aim to estimate the probability for a certain region to have more or less rain, or to be warmer or cooler than normal. According to Rienecker, NSIPP's goal is to estimate such probabilities 12 months in advance.

NSIPP's predictions start with the ocean general circulation model (OGCM), which is initialized by assimilating data to estimate the ocean state at the start of the forecast (see "Data Assimilation and Sources of Data"). El Niño events are not purely ocean phenomena. Because the ocean and atmosphere are linked dynamically in nature, NSIPP links the two in the subsequent forecasts. "You can think of the atmosphere as constantly adjusting to the ocean," says Suarez. "At the same time as that adjustment is done, the atmosphere forces the ocean." The atmosphere moves

A comparison of the structure of the equatorial Pacific thermocline under normal conditions (top) and during an El Niño (bottom). Temperatures—red represents 28°C and blue represents 8°C—are taken from NOAA's Tropical Atmosphere Observation moorings.



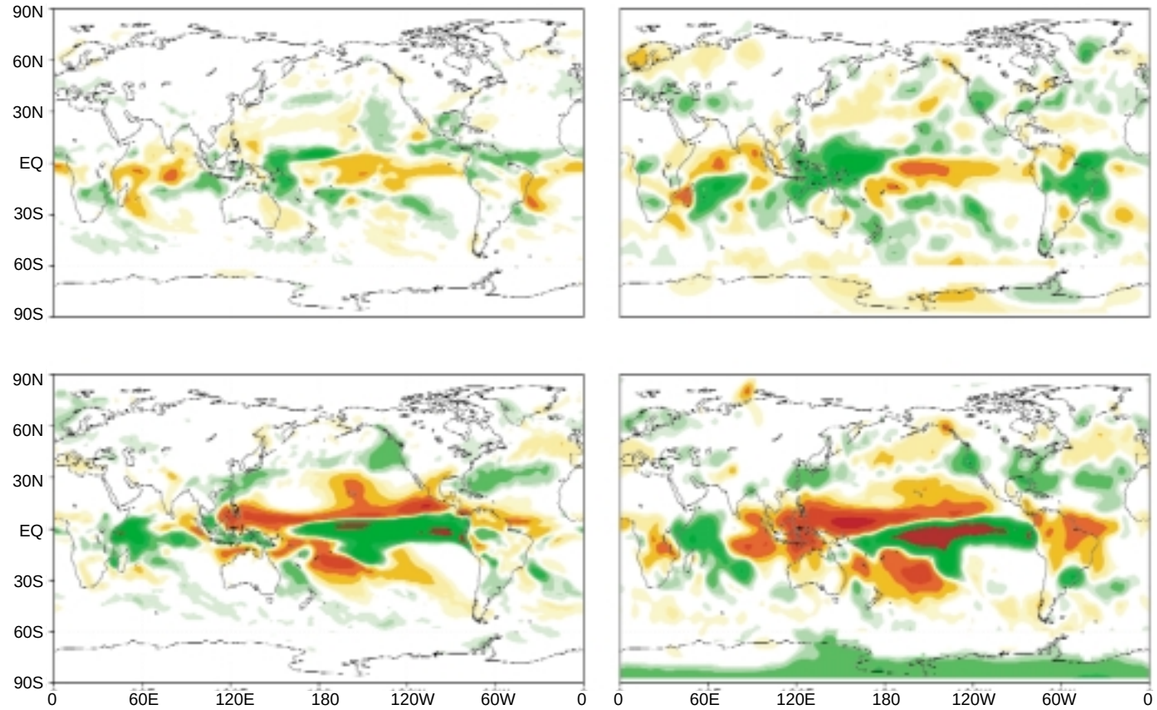
warm pools of water in the ocean, these warm pools create updrafts and affect the winds, and these changed winds go on to move more water. Once coupled, the OGCM integrates the equations of motion forward, predicting how the sea surface temperatures (SSTs) change in response to the surface winds and air–sea heat and moisture exchange as well as the ocean currents. Each day, the atmospheric global circulation model (AGCM) takes the SST forcing and produces a forecast that is then used to force the OGCM for another day. These coupled forecasts are usually conducted on 64 processors.

During FY99, NSIPP started to test the coupled model by doing hindcasts. Hindcasting is a retrospective forecast whereby scientists go back to, say, 1992, make a forecast for the 1993 El Niño, and see how well the model does by comparing the output to actual oceanic and meteorological data from 1993. Hindcasts are useful for improving regular forecasts because they allow NSIPP to determine in what areas the two models linked together simulate reality well, and in what areas they are deficient.

"After we find out what the El Niño is doing, then we find out what its consequences are elsewhere," says Suarez. This can be done with the coupled model. However, to estimate the probability that any forecast is reliable, NSIPP runs ensembles of atmosphere-only forecasts that are based on forecast SST from the coupled model. Typically, the NSIPP team uses the NCCS's T3E to run these forecast ensembles—nine or so, each with different initial atmospheric conditions—and then averages them. The tiny perturbations in initial conditions result in a

WHY PREDICT EL NIÑO?

The onset of an El Niño is detected easily by satellites and ocean buoys as the warmest water in the Pacific shifts eastward and normal trade winds relax. But knowing that an El Niño is happening is not as useful as being able to predict one months in advance. Given significant warning, governments and individuals might be able to mitigate the accompany ill effects of the strange weather that often accompanies El Niño events by setting aside resources for disaster relief from anticipated floods or droughts, taking action to protect beaches from erosion, and planting crops earlier or later than normal, or even planting different crops than usual.



Ensemble mean hindcasts from a coupled atmosphere–land–surface model using observed Reynolds SST data. Precipitation anomalies (in mm/day) averaged over January, February, and March (left), and observed Xie–Arkin anomalies (right). Upper panels are for 1989, a La Niña year, and lower panels are for 1998, an El Niño year. Such hindcasts show the predictability based on "perfect" knowledge of SST.

distribution of outcomes, some states being more probable than others, and the real climate perturbations will tend to fall somewhere in the middle. "That's the kind of forecast you can make," says Suarez.

The ensemble forecast is a computationally expensive calculation, typically using 144 processors, with each ensemble member running on its own 16-processor partition. Each calculation, while not pushing the T3E to its limits, is still a very large computing problem in absolute terms—it cannot be done on workstations. Processing speed is in the 5-GFLOPS range, and the output typically is 300 MB.

As part of their ongoing model development effort, Rienecker and Suarez also use ensembles to try to elucidate the background "error covariance" in their OGCM, or the relative contributions to the error of their forecasts from either the uncertainty of the parameters in the OGCM or the uncertainty in the surface forcing. For this, they use Monte Carlo simulations and an ensemble Kalman filter (EnKF). Both methods use slight perturbations in the surface forcing to generate ensembles. The EnKF is a statistical method that also involves running several forecasts with assimilated data that have been "perturbed" by adding random noise.

These computationally intensive activities are helped by NSIPP's close collaboration with the NCCS. "The benefit of the T3E," says Rienecker, "is that we have been able to do 96 different Monte Carlo simulations, which we were able to use to make estimates of the error covariance structures." The EnKF implementation is the first of its kind in a truly parallel environment for a state-of-the-art OGCM.

Another NSIPP activity is to conduct long simulations with the AGCM coupled to the OGCM and to a land surface model to investigate what limits the capa-

DATA ASSIMILATION AND SOURCES OF DATA

One of NSIPP's major accomplishments has been the completion of its ocean data assimilation system, which had been in testing for 2 years. Prior to that, cruder methods had been used to initialize NSIPP's ocean forecasts. The data assimilation system takes a multitude of oceanic observations from a variety of sources and produces an "initialization," which is a gridded field of ocean properties that is used as input for NSIPP's ocean model. The ocean model then takes the assimilated data and solves the equations of motion for the ocean, integrating forward for the length of the prediction.

For years, data assimilation of atmospheric weather data has enabled atmospheric models to make numerical weather predictions. Weather prediction looks at high-energy and small spatial and timescale variability in the atmosphere—specific storms and where they will blow in the coming days. Climate prediction, on the other hand, is concerned with less energetic, lower frequency, larger space scale, or longer timescale phenomena, and the ocean state is key to climate prediction.

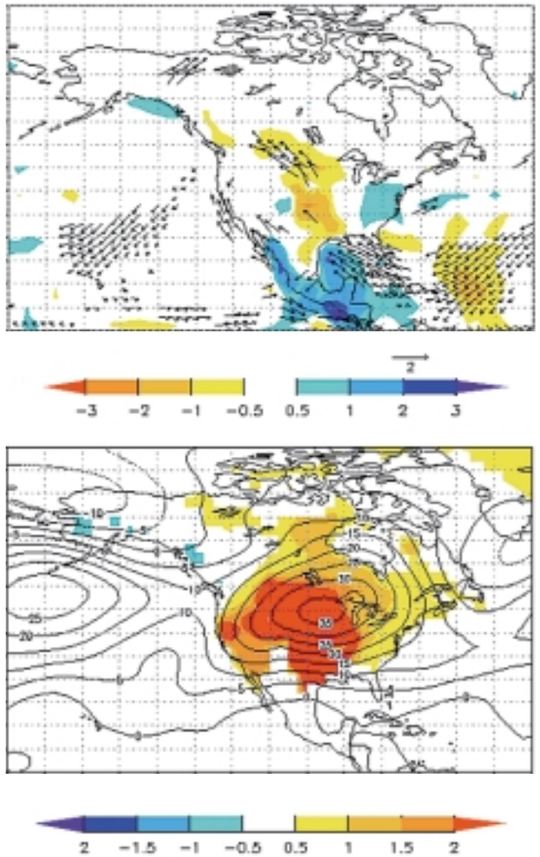
The key parameter in NSIPP's calculations is the wind stress at the ocean surface, which is measured by scatterometers. Scatterometers are specialized radar devices onboard satellites (e.g., QuickSCAT) that scatter microwaves off centimeter-long "capillary" waves, which are caused by wind blowing on the oceans. The speed and direction of winds over the oceans can be calculated.

Another crucial data source, sea surface height (SSH), is measured by laser altimetry satellites such as TOPEX/Poseidon. SSHs are critical because they detect the back-and-forth motions across the Pacific that are El Niño. The motions are most apparent on the thermocline, the region of the ocean that separates warm waters on the surface from cooler waters below. The vertical tilt of the thermocline is much less during an El Niño than during a normal year. NOAA's Tropical Atmosphere Observation (TAO) buoys, which extend along the equatorial Pacific, measure the thermocline motions directly. Other data sources include salinity, temperature at depth, and subsurface ocean currents, which can be measured only by instruments below the ocean surface.

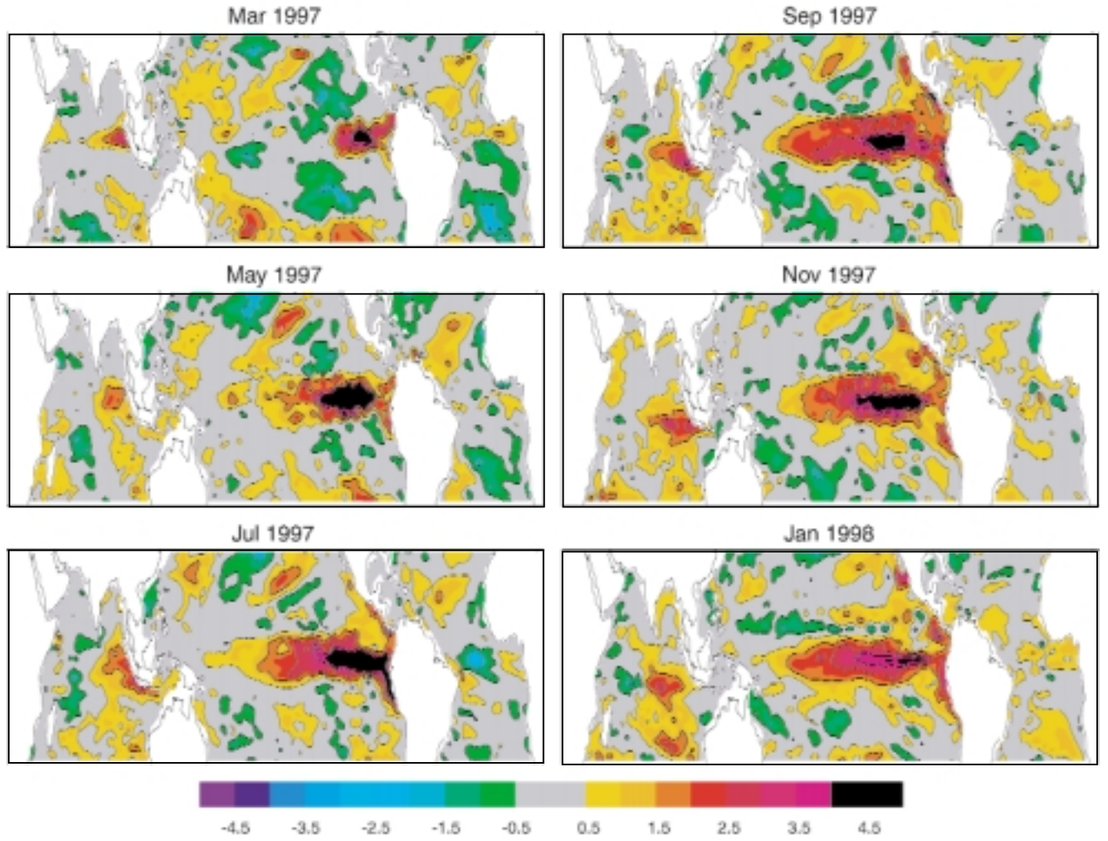
Waves on the thermocline produce slight variations on SSH, which can be measured with an altimeter. Buoys make more accurate thermocline measurements, but have the disadvantage of sparse coverage. "So our job at NASA is to see how we can get the most information out of those two systems," says Suarez. "Can we combine everything to get a better picture of what the thermocline is doing than from just the altimeter measurements or just the buoy measurements?"

SST is also an important data type, since it is the actual indicator that an El Niño is happening. SSTs are also crucial for initializing the model.

bility to predict El Niño many seasons in advance. Both the OGCM and the AGCM are written using object-oriented structures in Fortran 90 to facilitate passing messages between the two models. The incompatibilities of the individual models become apparent in a long coupled simulation, and the models can be tuned individually and improved incrementally. Using the T3E makes the process fast, which is useful for NSIPP. In the last year, a long coupled simulation was performed on the T3E. "We found deficiencies in our system as a result of doing that," says Suarez. Many of these deficiencies have now been corrected.



An example of 9-meter ensemble mean anomaly forecasts using the NCEP monthly mean forecast SST anomalies for June, July, and August. 850 mb wind anomalies plotted over precipitation anomalies (top), and 500 mb height anomaly contours with surface temperature anomalies (bottom).



The evolution of the forecast SST anomaly in centigrade. The forecast is conducted using NSIPP's coupled ocean–atmosphere–land surface model. The ocean is initialized using observed winds and SSTs and by assimilating subsurface temperature observations from TAO moorings across the equatorial Pacific. The forecast is initialized on February 1, 1997.

The coupled simulations have improved, and it is expected that this will affect the forecast skill of the coupled model.

One of the goals for the NSIPP ocean model is to eventually move to higher resolution in ocean grid size. The current version models the ocean on a grid that is 0.67° latitude by 1.25° longitude. NSIPP scientists aim to approach 0.33° for the ocean model and as fine as 0.33° globally for the coupled model. A 1° model produces four times the data as does a 2° model and is eight times more computationally intensive because of a further reduction of the time step. The advantage of increased resolution is that the calculations are more scalable—more processors can be used. A 1° model can be run on 256 processors and a 0.5° model can be run on 512. When running a simulation of 0.5° on 512 processors, it is possible to achieve speeds of 20 GFLOPS on the T3E. Conducting ensembles of simulations and forecasts at high resolution remains a computational challenge. The NCCS's T3E allows NSIPP to demonstrate the benefits to be gained by investment in such capabilities.

Some of the FY99 work at the NCCS has involved using larger ensembles in the forecasts to see how robust the small ensemble cases are.

LOW-LEVEL JETS

The Data Assimilation Office and Reanalysis

Data assimilation brings together atmospheric observations and atmospheric models—what we can measure of the atmosphere with how we expect it to behave. NASA's Data Assimilation Office (DAO) sponsors research projects in data reanalysis, which take several years of observational data and analyze them with a fixed assimilation system, to create an improved data set for use in atmospheric studies. Using NCCS computers, one group of NASA researchers employs reanalysis to examine the role of summertime low-level jet (LLJ) winds in regional seasonal climate.

Prevailing winds that blow strongly in a fixed direction within a vertically and horizontally confined region of the atmosphere are known as jets. Jets can dominate circulation and have an enormous impact on the weather in a region. Some jets are as famous as they are influential. The jet stream over North America, for instance, is the wind that blows eastward across the continent, bringing weather from the west coast and increasing the speed of airplanes flying to the east coast. The jet stream, while varying in intensity and location, is present in all seasons at the very high altitude of 200–300 millibars—more than 6 miles above Earth's surface.

LLJs, on the other hand, are confined to the bottom few thousand feet of the atmosphere. Because they are so low, LLJs are highly susceptible to the influences of conditions on the ground or in the ocean so that, for example, they may be present only at night and/or only during certain seasons.

Aerial photograph courtesy of NOAA.

The low-level jets are important in those seasons, though: The Somali jet, for instance, feeds the south Asian monsoon, and the Great Plains low-level jet (GPLLJ) is associated with summer thunderstorms in the Midwest.

North–south jets are particularly interesting to researchers because they bring the moisture and warm air of the Tropics to the subtropics and the extratropics. Because this warm, moist air originates in the Tropics, it is influenced by tropical climate signals, such as sea surface temperature (SST). The jets are one link between those tropical SSTs, which are somewhat predictable, and extratropical weather, which may not be. Understanding north–south LLJs could lead to useful climate and weather predictions over North America and Asia.

One of the DAO researchers trying to understand north–south LLJs is Siegfried Schubert, who, with his colleagues, uses reanalysis data and NCCS computer time to study the Somali jet and the GPLLJ. His ultimate goal is to use the jets and what he learns from them to make climate predictions for the regions they influence. "Predictions are incredibly important," says Schubert. Predicting what the atmosphere will do many months down the line is difficult, though, because of chaos. "Even if you had a perfect model of the atmosphere, errors would still grow in a forecast, because you would never be able to measure the initial conditions perfectly everywhere," says Schubert. "There will always be some error at the smallest scales, and, after a couple of weeks, the error will propagate to contaminate even the largest (planetary) scales."

This is why the two jets that Schubert and his colleagues study are so interesting. They know that the massive influx of moisture out of the Gulf of Mexico affects the weather over the Great Plains (see "The Great Plains Low-Level Jet"). By asking how that jet is affected by conditions in the Tropics and subtropics, some useful predictions might be made. Likewise with the south Asian monsoon: The exact mesoscale behavior—what happens on the scale of less than about 60 miles across—may be impossible to predict, but answers could be found for some less ambitious questions, such as when will the Asian monsoon start next year, and how intense will it be? (see "The South Asian Monsoon"). These are the sorts of questions that data reanalysis projects can address.

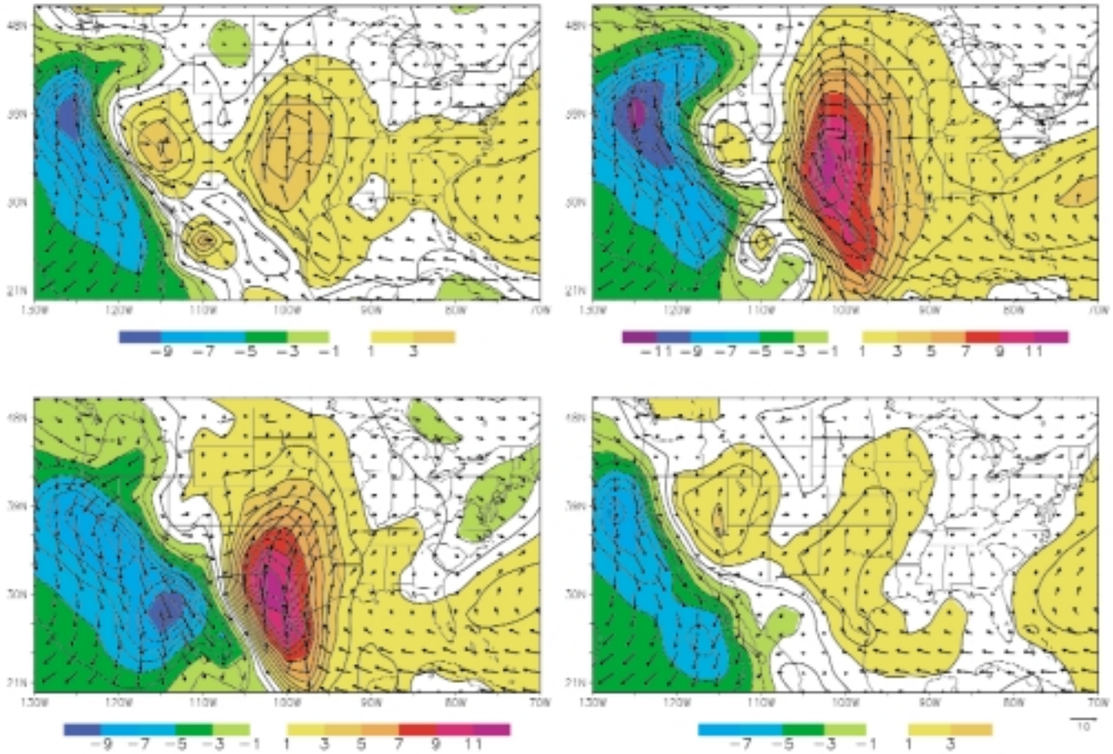
> NCCS support staff is very familiar with the DAO, its techniques, and its needs.

Schubert and his colleagues use a multiyear atmospheric data set generated with version 1 of the DAO's Goddard Earth Observing System (GEOS–1) data assimilation system for the period March 1980 through November 1995. Such an analysis of historical observations is often referred to as a reanalysis. The assimilated or reanalysis data show a structure for the jet that Schubert and his colleagues otherwise could not obtain. Observational data showing vertical wind motion, moisture and heat fluxes, and ground moisture are crude or nonexistent, but data assimilation takes whatever data do exist and interprets these variables. "The model is actually producing all that structure," says Mark Helfand, of the DAO.

As improvements are made to the atmosphere models and to the observing systems, the data for a given time period can be reanalyzed and an improved data set can be assimilated (see "How Data Assimilation Evolves").

"Will we ever have an observational system that will tell us everything we need to know about the climate and predict it?" asks Helfand. "I doubt it."

"But," Helfand adds, "I hope we don't. I hope nature is so unpredictable that we're always fascinated by it and never totally understand it. There are so many things going on in the climate. So many interactions. Nature is no machine."



Composite mean diurnal cycle of wind vectors May through August 1980–1994 at approximately 300 meters (m) above the surface. The shading represents the strength of the southerly wind. Units are ms^{-1}. Winds are shown at 6 p.m. (upper left), midnight (upper right), 6 a.m. (lower left), and noon (lower right), c.s.t.

THE DATA ASSIMILATION OFFICE

Data assimilation is a methodology that combines observations with a first guess from a model to produce global gridded and, in principle, optimal estimates or analyses of the entire Earth System. Data assimilation improves analyses beyond what observations alone provide, because it combines data types, propagates data to places where there are no observations, and provides estimates of difficult-to-observe diagnostic or forcing fields.

One such data assimilation system was developed and is managed by NASA's DAO. "Our mission," says DAO Head Robert Atlas, "is to advance the state of the art of data assimilation and to produce research-quality data sets through data assimilation." The DAO assimilates ovservations into a state of the art numerical model. The numerical model is based on current best understanding of the physics of the atmosphere and is parameterized to accurately reproduce this behavior. Data are brought in incrementally when they are available, and constrain the model's estimated values. A statistical analysis scheme corrects the first guesses to reflect the new data. "The data bring the model back closer to what the real atmosphere is doing," says Atlas.

Further error checking and mitigation are always required. For example, temperatures taken along cold or warm fronts may fluctuate greatly over a short distance. These observations may not translate well in a data assimilation system because the system will use the highly unstable temperatures to assign an average temperature to a large region—anywhere from a few dozen to several hundred miles, depending on the grid size of the model. Errors come from both the observations and the model, and both must be minimized.

The DAO's final product is a gridded, four-dimensional data set that has values for moisture, temperature, pressure, cloud parameters, and wind vectors over the grid points. The data sets are produced every 6 hours, though certain data sets are produced continuously in time. Scientists then use the data to study such things as Earth's angular momentum, surface fluxes, polar phenomena, moisture transport in the atmosphere, LLJs, sea ice drift, ocean stresses, and many other applications.

After the data set is complete, many validation studies are performed to ensure that the values of the atmospheric parameters produced by the DAO are accurate as compared to reliable measured values of those same parameters. Members of the DAO and outside researchers perform a certain amount of validation before the data sets are released to the community.

Then, as in any good science, outside, independent researchers use GEOS–1 in their own studies and make their own validations, informing the DAO when some values are called into question. "When we finished the GEOS–1 reanalysis," says Schubert, "we put the data out in the community and then had a workshop [in 1995] to get feedback."

HOW DATA ASSIMILATION EVOLVES

Atmospheric models attempt to best represent the atmosphere. While not perfect, these models are at least consistent. "They give us a picture of what the atmosphere could look like," says Helfand. Then, the weather observations provide discrete pictures of what certain parts of the atmosphere actually do look like in certain places at certain times.

However, weather data sources are many, varied, and scattered. There are about 250,000 data reports every day from satellite instruments, ground weather stations, weather balloons, ships, buoys, and aircraft sounders. Each report may include a variety of information, such as temperature, humidity, or pressure at a particular location and time, and the reports are unevenly distributed about the globe.

Making these disparate data useful to those who are trying to understand and predict the weather and climate is the goal of data assimilation. "Data assimilation is a way of taking the data that we measure and bringing it together with how we think the Earth System works," says Helfand.

New data types are added as technology improves and new instruments are added to the world's observing systems. Observing systems are always changing somewhat with each new satellite launched. Occasionally, a new type of observation arises, such as the scatterometer aboard the QuickSCAT satellite. To be useful to science, these new data types must also be added to the DAO's data assimilation system. "In the past, NASA produced data sets that sat on a tape for years with nobody using them," says Atlas. "Our goal is that, when a new data set becomes available, within weeks or months to be assimilating it."

New data types must have their error determined to minimize the error's effect on the assimilated product. "If you know what the errors are, you can correct them," says Atlas. "The better you can do this, then, the more accurate your product will be."

Another way that the DAO's assimilation system evolves is in resolution. The original GEOS–1 product was 2° latitude by 2.5° longitude, but since then the DAO has improved the resolution to 1° and 0.5° models. Limited-area 0.25° data sets have even been produced. "You need to improve the assimilation to the point where [the models] really do reproduce the atmosphere's behavior—particularly over a certain region," says Atlas.

In terms of predictability, the DAO's model is about halfway to the theoretical limit of weather prediction. "Seven days is around the limit of useful predictive skill now," says Atlas. "Model improvements, analysis techniques, and observations will give us the rest."

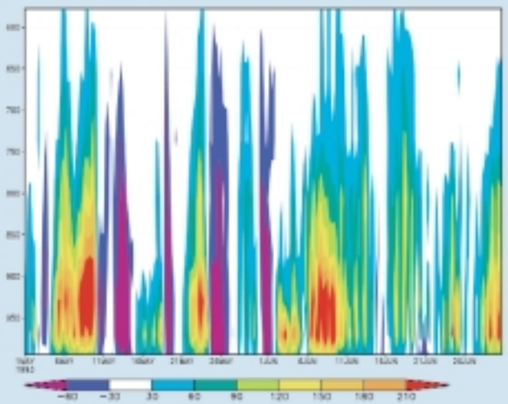
THE GREAT PLAINS LOW-LEVEL JET

On a warm summer night in the Great Plains, the wind may be barely perceptible at ground level. However, just a hundred or so meters above, and continuing up for a couple of thousand meters in the lower atmosphere, the wind blows strongly most nights. The wind turns clockwise over the Great Plains to complete its circulation over the eastern half of the continent. It eventually reaches the Chesapeake Bay and the northeastern U.S.

This warm, moist wind, originating in the Gulf of Mexico, carries as much as one third of the moisture coming into the central U.S. and, with it, rain. It is a warm-season phenomenon with a strong diurnal cycle, blowing primarily during the summer and at night. More than half of the thunderstorms that pour rain over the Great Plains States every summer occur between 6 p.m. and 6 a.m. This is unusual, because thunderstorms over land generally occur in the afternoon.

This nightly wind, the GPLLJ, is one of the LLJs of interest to Schubert and his colleagues at the DAO. Specifically, these researchers are investigating the dynamics of the transport of moisture by the LLJ as well as what happens to the moisture. The moisture flux is greatest at night because during the day the Sun warms the ground, which heats the air close to the ground, making the air become unstable, release energy, and undergo turbulent mixing with higher layers of air to smooth out the vertical profile of momentum. "The jet cannot be sustained in that environment," says Helfand.

At nighttime, the mixing ceases, and the GPLLJ accelerates over the Plains and rotates clockwise in what is referred to as inertial oscillation, which can be observed in wind vectors (speeds and directions).



Time series of northward moisture transport (υq) profile at 32°N, 97.5°W. Vertical scale is approximate pressure level assuming 1000-mb surface pressure. Units are ms^{-1} g kg^{-1}.

Because pressure differences and wind vectors are best observed through data assimilation (there are not enough actual observations), Schubert and Helfand use GEOS–1 data. GEOS–1 is useful, though crude, because of its coarse temporal and spatial resolution. "To have something that is useful in the end," says Schubert, "what you really want to do is to make high-resolution predictions over some region."

Schubert and Helfand would like to look directly at the impact of the GPLLJ on thunderstorms and other weather phenomena in the Great Plains in the summer. Such studies are not possible at the moment because summer thunderstorms occur on the mesoscale and beyond model resolution. What is possible, however, is for Schubert and Helfand to use the DAO's assimilation techniques and NCCS resources to create additional data to study the GPLLJ signal.

By performing reanalysis, Schubert can obtain assimilated data every 3 hours, instead of the usual 6, which is important since the jet has such a strong diurnal cycle, which is believed to be due to the inertial oscillation. Schubert can also focus solely on the region made up of the continental U.S. and the surrounding oceans and obtain data at a higher resolution over that area. Then he looks at how the GPLLJ varies from year to year and from month to month.

"We look in detail at the structure of the circulation at different timescales," says Schubert. "On which timescales does this jet show up, what is its periodicity, what is its variation from year to year, from one day to the next, from one week to the next?" Schubert and Helfand look at time mean fields, month-to-month fields, and average diurnal cycles averaged over a summer or over several years or more, asking how changes in boundary conditions—SST for instance—and other atmosphere, ocean, and land variables correlate with changes in the jet. They also ask what other diurnal, daily, weekly, seasonal, and interannual cycles are related to the GPLLJ.

Looking over 2–4-day timescales shows the effect of weather systems on the GPLLJ; these systems disrupt the GPLLJ as they move through. Averaging over a week or more makes it possible to examine the spatial and temporal extent of the jet—where and when it turns on and off. Averaging over a month or more also allows the seasonal dependence of the GPLLJ to be studied.

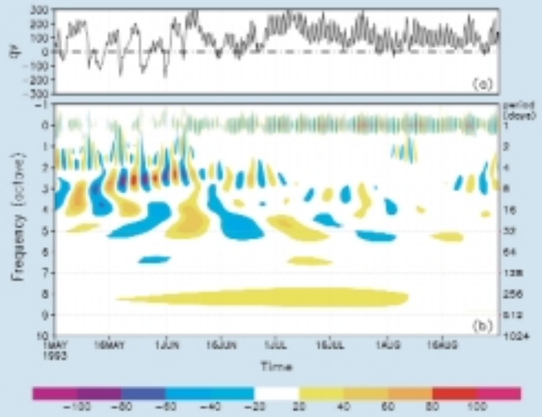
The GPLLJ has a strong seasonal dependence. The jet is barely detectable in April and sporadic in May because there are numerous passing 4–8-day low-pressure fronts enhancing the GPLLJ, and high-pressure fronts suppressing it. The GPLLJ slowly becomes more regular during the summer months.

Averaging over a season allows the calculation of the contribution of the GPLLJ to the total moisture budget for the Great Plains and demonstrates that the GPLLJ is the major contributor to the influx of moisture into the Plains. Averaging over an entire season is also useful for observing the influence of other large signals on the GPLLJ, such as the 500-year flood of 1993. During this flood, the jet was enhanced, which increased the influx of moisture and brought even more rain. "It was a very wet time," says Schubert.

Comparing monthly and seasonal averages from year to year—the drought of 1988 compared to the flood of 1993, for instance—can demonstrate how much the GPLLJ changes. More importantly, these sorts of studies can show how much of the annual monthly mean variation in the GPLLJ is random and how much of it depends on land hydrology, SST, or other factors that vary greatly from year to year.

The implications of these studies are that there may be a way to clarify the relationship between GPLLJ, precipitation, and boundary conditions at the start of spring. Identifying the influence of the boundary conditions might someday indicate how the average properties for a certain region would change with changing boundary conditions, allowing seasonal average rainfall to be predicted.

These studies are, according to Helfand, like any good science. "The more you learn, the more there is to understand."



Wavelet analysis of low-level northward moisture transport (υq) at 32°N, 97.5°W shown for May–August 1993. The top panel shows the time series of υq. The bottom panel shows the real part of the wavelet transform for each frequency. Units are $(ms^{-1}$ g $kg^{-1})^2$. The wavelet analysis was performed on a 9-yr period.

THE SOUTH ASIAN MONSOON

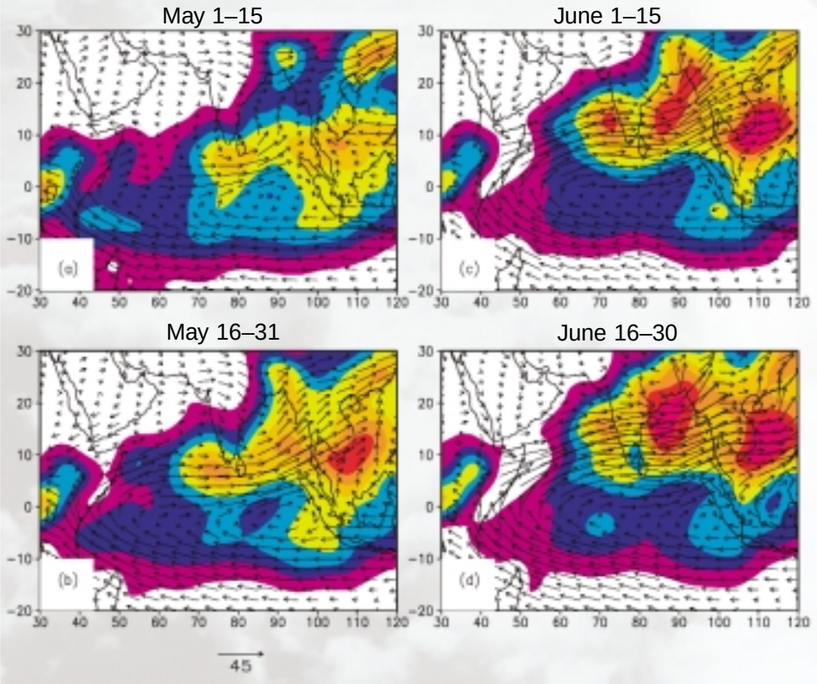
The heavy seasonal rains that arrive each year with the south Asian monsoon begin to fall over China and Indonesia in the early spring. As summer approaches, the rains reach the Bay of Bengal on the eastern side of India, usually within 1–2 weeks of the end of May. There the monsoon strengthens. It lasts around 40 days and becomes a spatially large and intense convective heat source over the region. By late summer, the monsoon has poured some of the heaviest rain in the world over India, Bangladesh, and the Bay of Bengal.

The South Asian monsoon is tied intimately to another summertime atmospheric phenomenon—an LLJ bringing low, warm, moisture-laden winds from the East Coast of Africa, known as the Somali jet.

Schubert and his colleagues are interested in the Asian summer monsoon both because it is a dramatic seasonal phenomenon that is a major player in general global atmospheric circulation and because it is the major source of rainfall throughout South Asia. It can also bring incredible devastation to the region. For example, flooding in Bangladesh kills hundreds, displaces thousands, and directly or indirectly affects hundreds of thousands of people. Damages sometimes reach into the hundreds of millions of dollars. Each year's monsoon is a little different—the amount of rainfall and the onset of the monsoon both vary, for example.

Each year is examined individually, and the monsoon is identified by the onset of the rains and the strong, circulating wind. The conditions in the atmosphere differ in the days and weeks before and after the monsoon, and are examined carefully. The wind fields indicate circulation patterns that also vary from year to year and are strongly forced by boundary conditions. This is a crucial connection, because some boundary conditions are predictable months in advance, which could mean that the circulation and, thus, the monsoon ultimately could be predictable. "The monsoon would be wonderful to be able to predict," says Schubert.

Predictions, however, will not come easily. A large project involving several international groups of climate modelers recently focussed on data from the 1997 and 1998 monsoons to examine the effect of the 1997 El Niño. "The thinking," says Schubert, "was that during a major El Niño event, the monsoon should be bad." However, as it turns out, the 1997 monsoon was near normal, which left Schubert and his colleagues wondering why.



Vertically integrated moisture flux and precipitation every 2 weeks for May and June averaged GEOS-1 climatology for 1980–1993. Units are g kg^{-1} ms^{-1} and mm/day^{-1}. The vertically integrated moisture flux is normalized by the surface pressure.

Tropical storm striking a sea wall, courtesy of NOAA.

"It didn't behave as we expected," he says.

"The monsoons are complicated and are affected by many atmospheric, ocean, and land events over short and long timescales," says Man Li Wu of the DAO. "For now, we are just trying to understand them."

So far, Schubert and Wu know that the Asian monsoon is affected by the Somali jet and that the Somali jet is affected by the enormous elevated land mass of the Himalayan mountain range and by interseasonal atmospheric, land, and ocean phenomena, such as SSTs in the Indian Ocean, ENSO events, snow cover in the mountains, and the amount of ground moisture at the start of the monsoon from the previous year's rainfall. What they do not yet know is to what extent the jet is affected by these phenomena.

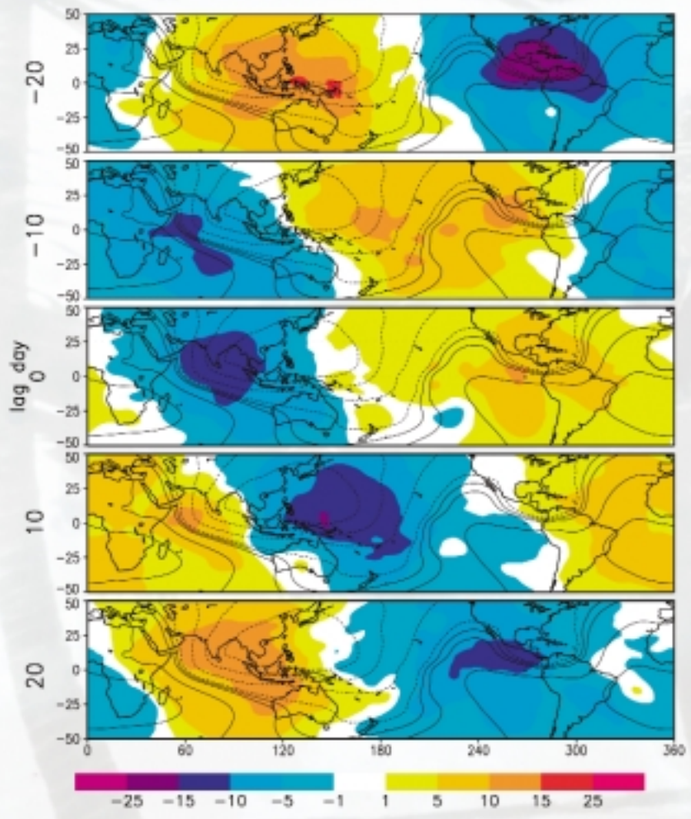
Schubert and Wu began to address this question in FY99. They ran an ensemble of 10 different simulations using the DAO's GCM. Each simulation had slightly different initial conditions but the same boundary conditions—SSTs for instance. "The idea was to see how much the boundary conditions constrain what happens," says Schubert. "How much of the variability is unpredictable versus how much of it can be accounted for by the sea surface temperatures and other things." A large spread in the predictions indicates a weak influence of SST, while a small spread indicates that the SSTs play an important role. They are still analyzing the data, using the NCCS supercomputers.

Something they have already discovered, though, is that the predictability changes with the season. It turns out that the SSTs are more prognostic of the Somali jet's behavior later in the summer. As the monsoon arrives in May and June, the winds are less predictable than in July and August, when the monsoon is already established.

They also know that the onset of the monsoon is affected by the global 40–50-day oscillation in pressure and winds first recognized by two scientists, Roland Madden and Paul Julian, then with the National Center for Atmospheric Research in Boulder, Colorado. The Madden and Julian oscillation (MJO) also affects the date of the break of the monsoon, which Schubert and Wu have demonstrated by comparing the dates of onset of the monsoon to the MJO signal seen in the reanalysis data.

In order to illustrate the quantitative extent of the MJO's effect, Schubert and Wu sought to decompose, or separate, the rainfall associated with the MJO from the total rainfall in the assimilated data. Wu compares this process to separating the yolks from the whites in an omelet. "Mother Nature is complicated," she says. "The signal she gives is scrambled."

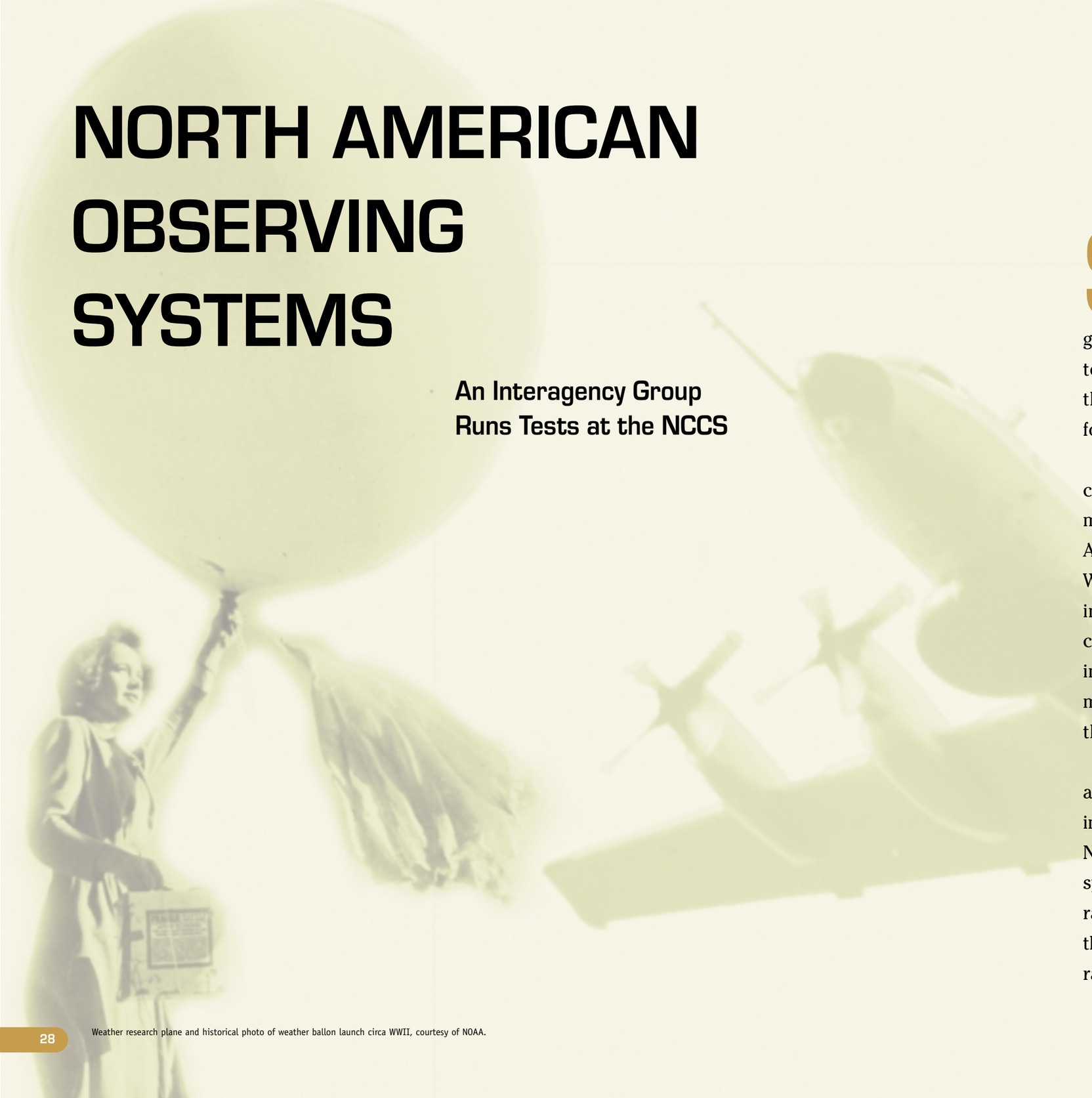
However, decomposition was made possible by using a new statistical method. This method, developed by Norden Huang of GSFC's Oceans and Ice Branch and known as the Hilbert-Huang Transform, sifts through data and extracts intrinsic oscillatory functions.



Composite average velocity potential for the MJO onset years (1980, 1982, 1983, 1984, 1986, and 1991). Lag 0 is the time of maximum rainfall over India based on the IMF 4 (MJO) index. Members of the composite have a 20–70-day filter applied. Units: 0.2×10^6 $m^2 s^{-1}$. The contours repeated in each panel are for the Jun–Aug mean (1980–93) velocity potential. Units: 1.0×10^6 $m^2 s^{-1}$.

NORTH AMERICAN OBSERVING SYSTEMS

An Interagency Group Runs Tests at the NCCS



Weather research plane and historical photo of weather ballon launch circa WWII, courtesy of NOAA.

Some 250,000 weather reports are collected by the National Weather Service (NWS) every day. Important measurements are taken by satellites, weather balloons, ground weather stations, airplanes, oceangoing ships, and tethered ocean buoys. Local or global weather models rely on these reports to provide the raw data used as initial conditions for the models to produce a weather prediction.

Because more accurate reports and higher quality initial conditions can produce better weather predictions, one of the missions of the National Oceanic and Atmospheric Administration (NOAA) is to bring about such improvements. While technological advances can lead to incremental improvements in measurements, implementing the technology can be expensive. Moreover, there is no guarantee that individual improvements will yield better overall forecasts. This makes decisions regarding new observing systems as risky as they are difficult.

To help North American governments and their agencies address upgrading weather observing systems, NOAA created the interagency North American Observing Systems (NAOS) Program. NAOS assesses the value of various combinations of observing systems to numerical weather prediction. In 1999, NAOS scientists ran a series of important experiments at the NCCS to determine the effect on weather prediction of removing some of the Nation's rawinsonde sites that provided possibly redundant data.

"The number of tests we had to do was daunting," says NOAA's Tom Schlatter, who chaired the group that did the study. "It would have taken years to get through them, and it's only because NASA stepped forward and said 'We'll help you' that we've been able to do anything at all."

Numerical weather prediction requires developing mathematical models that simulate the behavior of the atmosphere, solving the models' equations of motion on computers, and forecasting such variables as where the wind will blow and how air pressure, temperature, and moisture will change. The equations in these numerical models (based on Newton's second law of motion—acceleration is directly proportional to applied force, or F=ma) demand initial conditions that are based on accurate observations of the atmosphere's behavior and enable the models to reproduce that behavior.

Data assimilation systems make the connection between atmospheric observations and numerical prediction models; they initialize the models using weather data from a variety of sources (see "Data Assimilation in Numerical Weather Prediction"). These sources evolve as new instruments become available. Improving weather observing systems—and thus initializations—is one good way of improving weather forecasting. Ensuring that new weather observing systems actually improve forecasts is one of the most crucial issues in their design and is the mission of NAOS (see "Testing Atmospheric Observing Systems").

NAOS is composed of two separate working groups: The Systems Design Working Group, which analyzes the operations of observing systems, and the Test and Evaluation Working Group (T&EWG), which assesses the effect of new and proposed observing systems and configurations on weather forecasting. The T&EWG forms and tests hypotheses by running weather predictions over the same period of time while manipulating the observing system, adding or deleting components, and then evaluating the results. The group then makes its recommendation to the NAOS Council, which, in turn, makes the ultimate recommendation to all of the U.S. Government agencies.

The T&EWG used the NCCS's computers to address whether some of the rawinsonde sites could be eliminated without diminishing the accuracy of weather forecasts. Rawinsondes are weather balloons carrying

Map of rawinsonde sites. Sites designated with red triangles are part of a Global Climate Observing Network. Sites designated with bright yellow rings are near airports and were proposed for removal.

instrument packages that measure temperature, humidity, and pressure continuously in flight and transmit this information to a weather station, which also tracks the flight path of the balloon for information on wind speed and direction (see "A Typical Launch"). A network of 102 rawinsonde launching stations across the U.S. and surrounding territories gathers data twice daily.

Some sites are close to airports, and aircraft often take weather measurements in what is known as the Aircraft Communications, Addressing, and Reporting System (ACARS). Commercial aircraft file ACARS wind and temperature reports—about 50,000 a day over the continental U.S. and its coastal areas—and approximately 5–10 percent of these are recorded during takeoff or landing. "Descending or ascending aircraft," says Stephen Lord of NOAA, "produce a vertical sounding of temperature and moisture in the atmosphere." In theory, these 2,500 to 5,000 ACARS soundings could be used as a cheaper alternative to rawinsondes.

Experiments in 1996 demonstrated that reducing the rawinsonde network by half would substantially degrade the ability of numerical models to make accurate weather predictions in the 0–4-day range. But could some of the ACARS sites replace the nearby rawinsonde sites? "Could we then dispense with these sites and substitute the information derived from departing and arriving aircraft?" asks Schlatter. If the U.S. could stop paying for 14 sites, for example, it would save about $1.5 million a year.

This hypothesis is by no means self-evident, because ACARS data are somewhat inferior to rawinsonde data. Many ascent/descent soundings by aircraft lack the vertical resolution of rawinsondes. Moreover, aircraft did not begin measuring humidity until recently; only a handful of aircraft do so, even now. Rawinsondes also have the advantage that they rise to about 72,000 feet, much higher than an airplane's cruising altitude of about 30,000 feet.

Nevertheless, the T&EWG asked whether some of these rawinsonde sites could be eliminated and established several criteria for selecting the most likely candidates. Sites would have to be close enough to an airport that the climate would be roughly the same, atmospheric soundings would have to be taken during at least 20 ascents and descents per week, the distribution of the remaining sites would have to be roughly uniform, and no sites deemed important for the climatological record would be removed. In 1999, NAOS identified 14 sites that met these criteria and performed experiments to determine whether their removal would affect weather predictions.

The NAOS experiments were designed to observe a broad spectrum of weather conditions for long periods using several different operational weather prediction models. Everything normally done in a routine numerical weather prediction must be done for the experiments, requiring a large number of repeated and lengthy calculations run over long periods of time.

The T&EWG wanted to use well-tested weather prediction models to avoid model-dependent systematic error. They chose three operational models used by the NWS for daily predictions on the basis that these models had a lengthy performance record and the benefits of long-term verification and extensive feedback from forecasters throughout North America. They also used the data assimilation systems developed for each of these models.

A global model, the Global Spectral Model (GSM), depicts the atmosphere over Earth's entire surface from the ground to the stratosphere. The GSM has a resolution of about 1° latitude and longitude and 28 vertical levels. Two regional models, the Rapid Update Cycle (RUC) and the Eta, cover the continental U.S. and portions of Mexico and Canada. The RUC is a

TESTING ATMOSPHERIC OBSERVING SYSTEMS

NOAA created the interagency, intergovernmental NAOS Program in the mid-1990s to advise the U.S. and other governments about the most effective ways to spend their resources to produce useful atmospheric observations. NAOS provides recommendations to NOAA leaders and is guided by representatives from U.S., Canadian, and Mexican governmental agencies and universities.

Because any NAOS Program decision could affect its mission, NASA is involved. "Our concern [is] the observing systems not controlled by NASA," explains Robert Atlas, head of Goddard's Data Assimilation Office. "We want to make sure that if there are any changes to the observing system, they not adversely affect our ability to monitor climate or perform research."

NAOS experiments measure the relative contributions of different components of observing systems to numerical weather prediction. "We're trying to understand the value of observing systems to forecasting the weather," says NOAA's Stephen Lord. Some of these tests are simulations on new or proposed weather systems, such as Doppler wind lidar, and some are sensitivity tests to determine the effect of existing systems on weather prediction.

Sensitivity tests try to answer such questions as which are the best atmospheric observations to make, where, how, and how often. And, says Tom Schlatter, "What is the best way to put information we can glean from observations into a model and tell the model, 'This is your starting point. This is where you begin to produce forecasts.' "

Working groups within NAOS produce reports recommending directions. NAOS's TEWG examines new or proposed observing systems and helps the Government decide in a logical, scientific way whether they should be implemented.

"There have been many times in the past that very costly observing systems have been put aboard satellites and sent into space without any assurance that the information—the raw information that came back—would lead to improved forecasts," says Schlatter. "This is one of the things that we are trying to correct. We would like to know—before committing $100 million—whether that money would make a difference."

DATA ASSIMILATION IN NUMERICAL WEATHER PREDICTION

The accuracy of numerical weather predictions depends on using atmospheric observations to determine the weather conditions at the start of the prediction. Without this "initialization," no numerical model could correctly predict the weather.

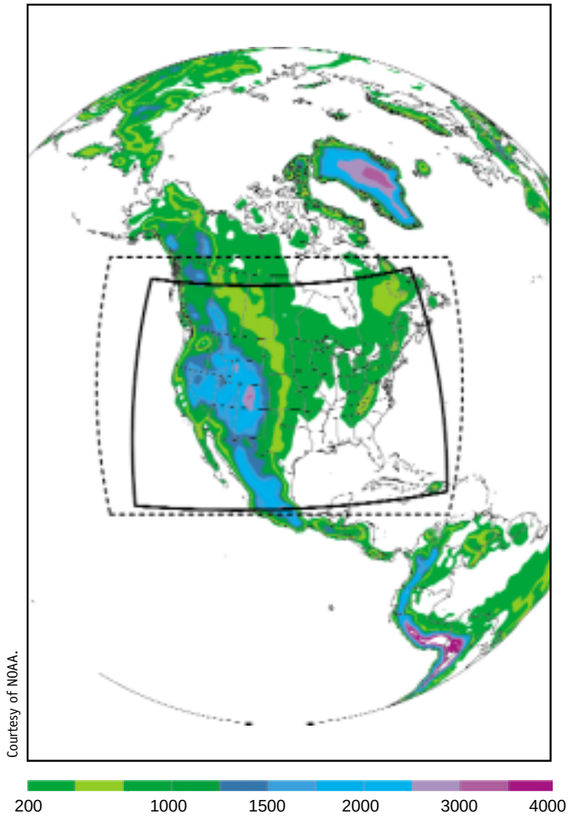
Weather data sources are many, varied, and scattered. There are about 250,000 weather reports every day from a combination of satellite imagery equipment, ground weather stations, weather balloons, oceanborne vessels, buoys, and aircraft sounders. Each report may include a variety of information, such as temperature, humidity, or pressure at a particular location and time, and the reports are unevenly distributed about the globe. Making all of these disparate data useful to meteorologists who are trying to predict the weather is the challenge of data assimilation.

"Data assimilation is a way of using observations to estimate everything that a weather prediction model needs to know to make a forecast," says Schlatter. To predict what the atmosphere will do, the data must be collected, condensed, and somehow converted into an orderly form.

Data assimilation systems specify atmospheric conditions on a regularly spaced grid of points that numerical weather prediction models use as initial conditions. All of the reports from the various weather instruments are used to estimate average properties of the atmosphere at regularly spaced intervals all over the globe, from the surface to the stratosphere, resulting in an ordered matrix that can be read easily into a computer.

"The beauty of a data assimilation system," says Lord, "is that it can take information that is produced at one point and propagate it into another part of the atmosphere that doesn't have an observation."

mesoscale analysis/forecast system that updates weather conditions hourly and produces short-range (12-hour) forecasts over the contiguous 48 United States and adjacent areas every 3 hours. The Eta model has been used by the National Centers for Environmental Prediction (NCEP) since 1993 to provide early forecast guidance over North America. NCEP runs the Eta model twice daily, producing a 60-hour forecast at a resolution of 32 km with 38 vertical levels. Each model uses its own data assimilation system to initialize the conditions when the NWS starts to run them.



Courtesy of NOAA.

Global land elevation, in meters, used for the GSM. The area of the Eta model is shown with the dashed line, and the area of the RUC model is shown with the solid line.

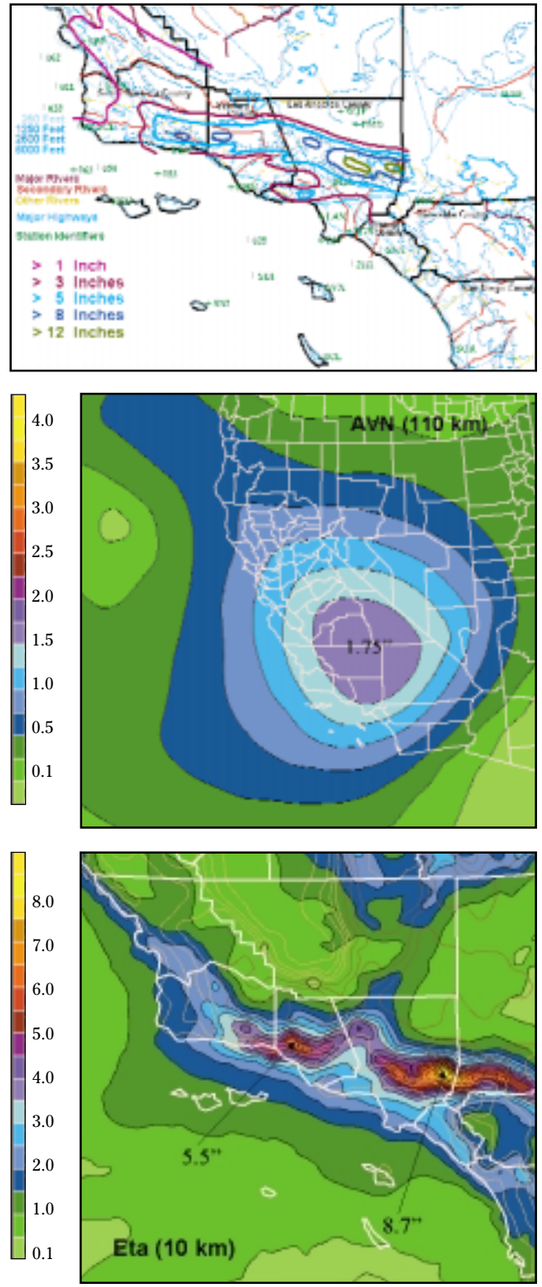
The plan was to run the experiments several times throughout the year over many different meteorological conditions to discover any seasonal variations in the results. In the first experiment, which was completed in 1999 using NCCS resources, the models were initialized with all operationally available data; the global model was used to predict the weather from December 1997 through February 1998, and the two regional models were used over a shorter period—January 25 to February 28, 1998—which was less computationally demanding. NAOS selected these months because they coincided with an active El Niño winter marked by large rainstorms in California and other significant weather events in January and February.

Each model was run three times: A control using the full operational data—which included rawinsonde and ACARS data—and two test cases. One test case excluded the rawinsonde data from the 14 sites and the other excluded both the rawinsonde data and the nearby ACARS data.

The results showed that there is very little loss of forecast skill with the 14 sites removed if the nearby ACARS data are retained. The most notable exception was that the loss of the rawinsonde moisture soundings contributed to clearly inferior moisture analyses in all three models. The early hours of the forecast suffered as a result, but, within 12 hours, this handicap had essentially disappeared. Otherwise, there were no differences between the simulations with all the data and the simulations with both the rawinsonde data and the ACARS data removed when the GSM was tested. There were only slight and statistically insignificant differences when both data types were removed with the Eta model. The RUC model prediction was also slightly less accurate when both data types were removed, particularly for the moisture field. "Not to the extent that a forecaster would have changed his public forecast," says Schlatter.

The NAOS council decided to end the tests at that point because the preliminary results were not convincing enough to support removal of the 14 rawinsondes. NOAA solicited advice from the climate community and forecasters in the field, both groups making strong cases for keeping the rawinsonde sites. Climatologists use rawinsonde data for their decades-long studies of global weather trends, since some of the sites have been launching balloons continuously since World War II and continue to contribute to valuable long-term data sets. Weather forecasters use the balloon launches for their daily weather predictions and rely heavily on the moisture soundings. "Our recommendation will be against any reduction in the rawinsonde network," says Schlatter.

The apparent inability of prediction models to "remember" humidity information supplied at the beginning of a forecast is a well-recognized problem. "No one should conclude that humidity information is, therefore, unimportant," says Schlatter. Part of the problem is that the humidity field is undersampled. Few ACARS aircraft collect moisture data, but the number should increase in the future. Currently, however,

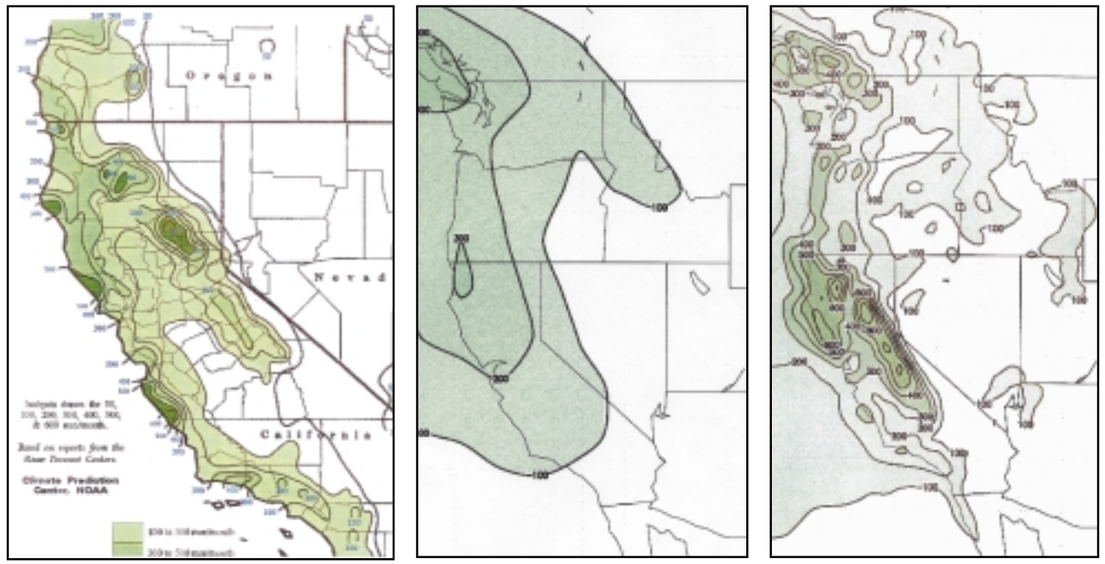


A comparison of observed rainfall, a GSM model, and an Eta model, shows the observed rainfall in inches (top), the output of the low-resolution GSM model (middle), and the output of the high-resolution Eta model (bottom).

rawinsondes are the only well-distributed source of three-dimensional moisture observations available. Also, humidity is the most difficult parameter to analyze and predict, and current research is focused on this problem. Moisture concentrations can vary significantly between rawinsonde sites and between the points of the computational grids employed in today's models. Another part of the problem is that accurate moisture and wind observations are required at the initial time before a model can generate correct vertical motions, clouds, and precipitation. More sophisticated data assimilation techniques that successfully couple the wind and moisture analyses would be helpful. Finally, the model's treatment of humidity can still be improved.

Schlatter predicts that the moisture information will become more important as new, more sophisticated models are introduced. "Overall," he says, "we learned a lot. And we couldn't have done it without the NASA computers."

NASA provided assistance through the NCCS, which arranged for the T&EWG to carry out experiments on one of the Cray J932se supercomputers. "Without help from NASA, these experiments would not be done yet," says Schlatter.



Examples of observed precipitation totals and model predictions in mm/month. (left) rainfall reported by the River Forecast Centers, (middle) low-resolution prediction of the GSM, and (right) high-resolution prediction of the RUC model.

A TYPICAL LAUNCH

At the Vaisala rawinsonde factory in New England, the Model RS80 rawinsondes used by the NWS are assembled. Inside these light-blue, cigar-box-sized cardboard and Styrofoam packages are thermometers, hygrometers, and barometers that measure temperature, humidity, and pressure.

The rawinsondes are shipped to the National Weather Center in Kansas City, Missouri, where each one is assigned a unique call number before being shipped to an upper air weather station. There is a network of more than 100 such stations across the United States and Canada, and many hundreds worldwide.

At 0000 and 1200 GMT, NWS employees attach a battery to the radio transmitter in the rawinsonde, activate the instruments, attach the rawinsonde to a weather balloon, and release the package into the air.

The radio transmitters send temperature, pressure, and humidity data back to the launching station, and the balloons are tracked with radar to get wind directions and speeds as they rise to the lower stratosphere over the course of several hours, toward an ultimate height of about 72,000 ft. There, the pressure around the balloon is so low that the helium inside forces the balloon to pop. A parachute attached to the rawinsonde allows the package to float back to Earth. An envelope addressed to the NWS center in Kansas City is attached in case the package is found—about half eventually are returned. Often, one rawinsonde may be flown several times.

Launch of a rawinsonde balloon.

CLIMATE PREDICTION SEES FUTURE DESPITE CHAOS

Researchers Outside NASA Use NCCS Resources for Studies

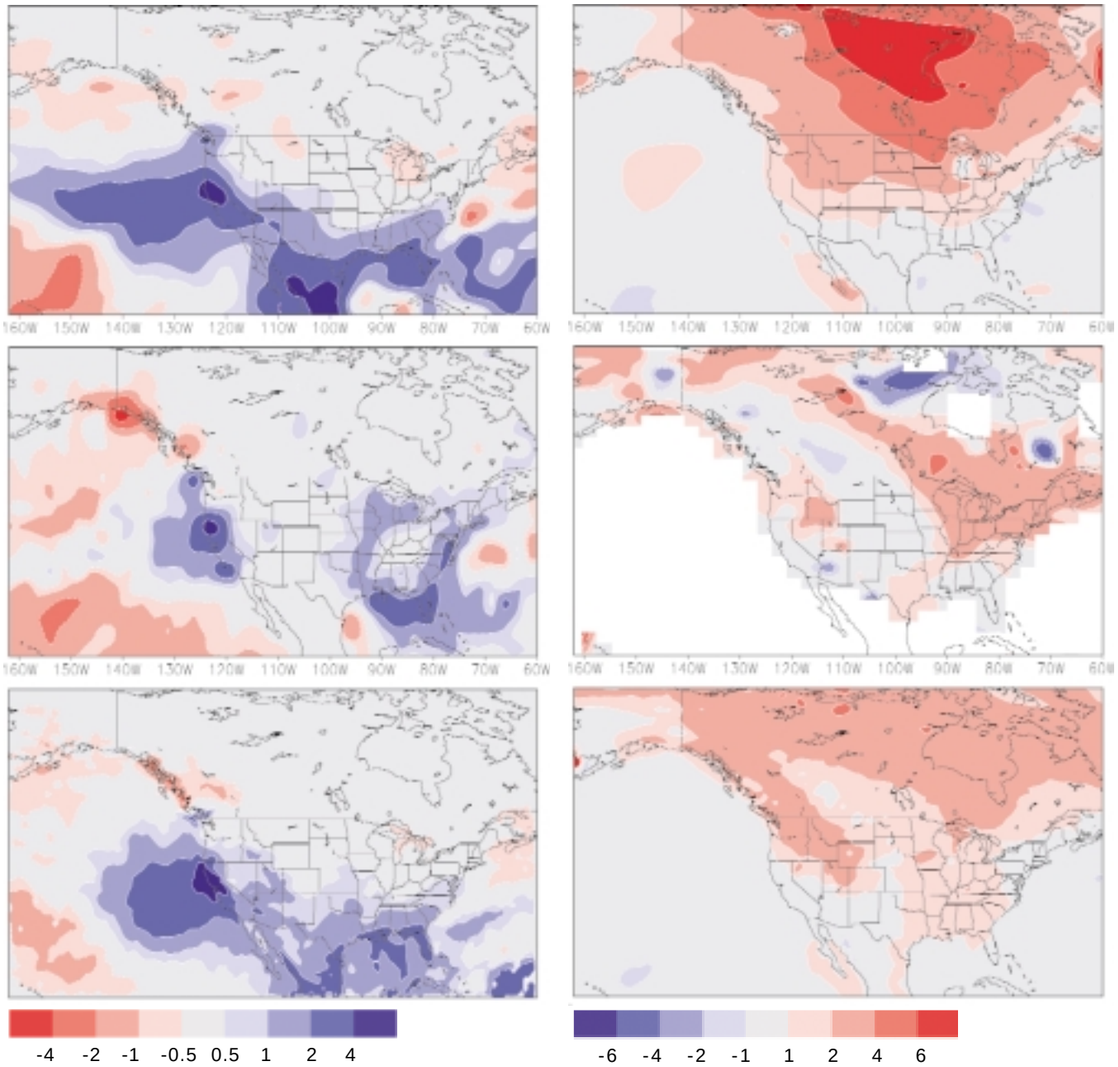
The air on this mostly sunny January day is crisp and the wind is blustery. The morning's National Weather Service 6-hour forecast had accurately predicted these conditions for the Baltimore-Washington area and the 2–3 day extended outlook was almost perfect. The previous week, the National Center for Environmental Prediction's (NCEP) 6–10 day temperature and precipitation outlook for the general trends for the region was correct as well.

However, no forecast could have predicted specific details about this day. It is 28.5°F in the sunshine, bright enough for dark sunglasses, and windy enough to blow off a hat. Such details are impossible to foresee with any accuracy and are outside the scope of routine weather prediction. Equally difficult is accurately forecasting weather beyond about 2 weeks. Jagadish Shukla, of George Mason University and the Center for Ocean–Land–Atmosphere Studies (COLA), is exploring the possibilities beyond these limits.

"What is predictable beyond the weather?" asks Shukla.

Weather forecasts predict specific conditions of the atmosphere for the near future, up to a few days, or the average properties of the atmosphere over longer periods of time. Predictions are always averaged over a specified region, perhaps the size of a small city. Weather forecasts use both these predictions and information from radar stations and weather balloons.

Lorenz strange attractors. Courtesy of Katharine Gurski, USRA.



Predicted precipitation anomalies for January–March 1998 from COLA's AGCM alone (top), Xie–Arkin rain gauge data from the NWS (middle), and nested Eta model (bottom).

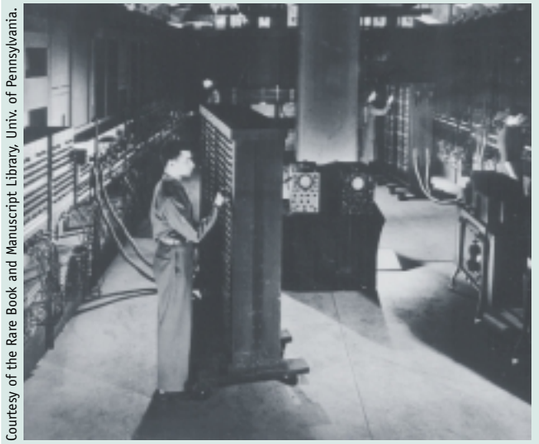
Predicted surface temperature anomalies for January–March 1998 from COLA's AGCM alone (top), Climate Analysis Monitoring Systems (CAMS) data from the NWS (middle), and nested Eta model (bottom).

Predicting the weather has inherent limitations because the atmosphere is a chaotic system (see "Chaos…."). According to chaos theory, the maximum lead time for a useful numerical weather prediction is about 2 weeks. Meteorology is still approaching this limit, and the most powerful weather forecasts today are useful to perhaps 6–10 days.

However, the atmosphere has certain longer-term fluctuations that may be predictable far beyond the 2-week limitation of numerical weather prediction. For example, what will the seasonal mean temperature be in the South next summer? Will the east coast have another drought? How much rainfall should Californians expect for the spring? Such questions are too far into the future to determine with numerical weather prediction.

Shukla and other scientists hope to answer these questions through numerical climate prediction, which uses the same basic methodology as numerical weather prediction. Forecasts of climate differ from those of weather, though, in that they are broader in scope. Broader can mean asking for the average rainfall over several months instead of breaking the rainfall down by day, or it could mean looking at a large region within the U.S. instead of a particular city or State. Numerical climate prediction potentially can forecast conditions months or seasons in advance.

CHAOS AND THE NONLINEAR DYNAMIC SYSTEM WE CALL THE WEATHER



Courtesy of the Rare Book and Manuscript Library, Univ. of Pennsylvania.

View of the ENIAC, which was equipped with 17,468 vacuum tubes, filled a 30-by-50-foot room, and weighed 30 tons.

Ever since the first computer, ENIAC, and its progeny machines, such as MANIAC, were developed at the University of Pennsylvania and Princeton University in the 1940s and 1950s, one of the toughest problems that scientists have tried to solve using the fastest and best computers is also one of the oldest: What will the weather be later today, tomorrow, next week, or next month?

Weather models use observations and numerical methods to represent the atmosphere. Computers first translate empirical weather observations into initial atmospheric conditions, and then apply to those initial conditions equations that reflect the physical behavior of the atmosphere, integrating the numerical forms of those equations ahead in time to try to predict future events.

Of course, the fact that weather predictions are so often wrong demonstrates that there are problems with the models, the initial conditions, or both. A perfect prediction would necessitate perfect knowledge of the initial conditions and a model that perfectly simulates the atmosphere with no error.

Error in numerical weather prediction comes from two main sources. One is the inherent inaccuracies in the models scientists use to describe the atmosphere. Basically, these models are equations that take some set of observed or mathematically derived inputs (the initial conditions), integrate the equations ahead in time, and produce predictions based on the results. All models make certain assumptions about how the atmosphere behaves and incorporate these assumptions into the models in the form of parameters—fixed numerical quantities that fit the equations to observations. The fit may not be exact, and the observations have some degree of error, so the parameterizations and, thus, the models all have some degree of inaccuracy. Different parameters give different results, and so do different models using different parameters.

The other inherent inaccuracy in weather predictions is much more profound. Chaos theory posits that any two predictive forecasts produced using identical models but with ever so slightly different initial conditions will, in time, produce drastically different results. Since there is no way to know exactly what the weather conditions are like at any time given the limitations of observing equipment, slight, even microscopic, errors in measurements will turn into very inaccurate predictions within a short period of time.

The propagation of error is determined by chaotic dynamics, and the theoretical limit to accurate weather prediction is about 2 weeks. After that, small errors in the estimates of the initial conditions will balloon into one gigantic wrong forecast. "You cannot fight chaos," says Shukla. "You cannot control it."

This phenomenon, following the work of Edward N. Lorenz of MIT, is known as the butterfly effect: The wind produced from the flap of a butterfly's wings on one side of the globe will affect the weather a few weeks later on the other side of the globe. Since there is no way for weather observations to observe such nuances, weather is predictable only up to a point. Meteorologists and climate scientists have generally come to accept this essential limitation of numerical weather prediction. "That's granted," says Shukla. "That's chaos."

The NCCS contributes to climate prediction research efforts by providing the supercomputer and mass storage resources that allow extensive calculations.

Climatologists hope that, by asking broader questions, they can push the limits of weather predictability and still find useful answers. Shukla's aim is to use numerical climate models to predict with greater than 50 percent accuracy such things as mean seasonal rainfall and mean temperature anomalies.

Numerical climate prediction uses mathematical models—sets of equations based on physical laws and certain assumptions about the behavior of the atmosphere. Like weather prediction, climate prediction uses observed data to produce the initial conditions—the process of data assimilation—and integrates the numerical atmospheric models forward in time. The end product, or prediction, is simply a projection of the final condition of the atmosphere after some number of time steps. The projection could be several days in weather prediction or several months in climate prediction.

Shukla and his group use a three-tier process to make these predictions. A low-resolution, coupled ocean–atmosphere model produces sea surface temperature (SST) data that are input to a medium-resolution atmospheric general circulation model (AGCM). This AGCM, based on NCEP's operational medium-range weather forecast model, incorporates—nests—another general circulation model in the region over North America. The nested regional model, called Eta, is a high-resolution mesoscale model that explicitly accounts for land topography. The more time-consuming steps in this process are performed on charney, the NCCS's Cray J932se supercomputer.

The most important input that Shukla uses is SST in the Tropics (see "Why Are the Tropics So Important?"), although he also uses SST data in latitudes outside the Tropics, soil moisture, and the previous year's snowfall in the regions of interest over North America.

SST data are taken from satellite data, buoys, and shipboard soundings and fed into a low-resolution, coupled land–ocean model to generate a regular SST gridded product. Then, these gridded SST data and data assimilated from weather-balloon, remote-sensing, and satellite observations are used to initialize the global model with its nested, high-resolution Eta model to forecast atmospheric conditions over the North American region. Ground moisture and snowfall data are incorporated in the regional model from operational observations.

What allows the long-term predictability in the first place is that SST changes are slow and are not sensitively dependent on initial conditions. Sensitive dependence is a term mathematicians use to describe what happens in a chaotic system such as the atmosphere. The atmosphere's chaotic nature wreaks havoc with the error estimates in measured quantities used to initialize numerical models of the atmosphere. The equations used to model the atmosphere are nonlinear and have a root mean square error doubling every 2 days. This means that, in a weather prediction, the error in estimated initial conditions at the beginning of the prediction quickly propagates the uncertainty of the prediction, rendering it of no use. But SSTs, while also dynamically forced by the atmosphere, are not sensitively dependent on the initial conditions of the atmosphere.

"We have now demonstrated the seasonal predictability of regional climate over North America," says Shukla, "and that's a new thing."

Shukla demonstrated this in 1998 by carrying out separate simulations using the same SST data with different initial atmospheric conditions observed on different days. The separate simulations converged to similar approximate average rainfall anomalies after the model integration, which were close to the observed anomaly. Thus, sensitive dependence did not play a role in the prediction.

The computational requirements for Shukla's studies are vast. SST and other data are first assimilated and then assigned to 2.8° longitude by 1.8° latitude grid points and 20 vertical levels over the surface of the globe. The model is then integrated forward in time in steps of 10 minutes for up to 100 days. With 10 equations for every gridpoint, the resulting calculation can be as large as several hundred billion calculations, each consisting of many hundred floating point operations.

"That's why we needed the NASA computers," says Shukla. "We couldn't have done it otherwise. Those supercomputers have been able to help us analyze huge amounts of data and to define the initial conditions much better." To be able to predict these average temperature and rainfall anomalies in advance with accuracy is impressive, considering the theoretical 14-day limit for predictability of the weather.

COLA is asking what other sources of information might reveal long-term signals in weather patterns—for example, ground moisture. Further studies will continue to draw on NCCS resources, especially as Shukla's group at COLA attempts to produce finer and finer maps of regional climate predictions, since each doubling of the resolution increases the computer resource demands 16-fold.

But, says Shukla, computational demands are not the only limiting factor. The most important issue is how the climate is modeled. A dynamic system such as the atmosphere demands a fully dynamical model. Fully dynamic weather prediction models have been in routine use for decades, and Shukla predicts that science is on the verge of developing a dynamical model that will produce climate predictions routinely. But before this can occur, Shukla says, there must be improvements in the way that models capture the current climate.

"We don't have good models. We must build good models. People used to say, 'We need more data' or 'We need faster computers,' but are these the only things limiting climate predictions today? No! What is most affecting predictions today is the limitations of the models," says Shukla.

WHY ARE THE TROPICS SO IMPORTANT?

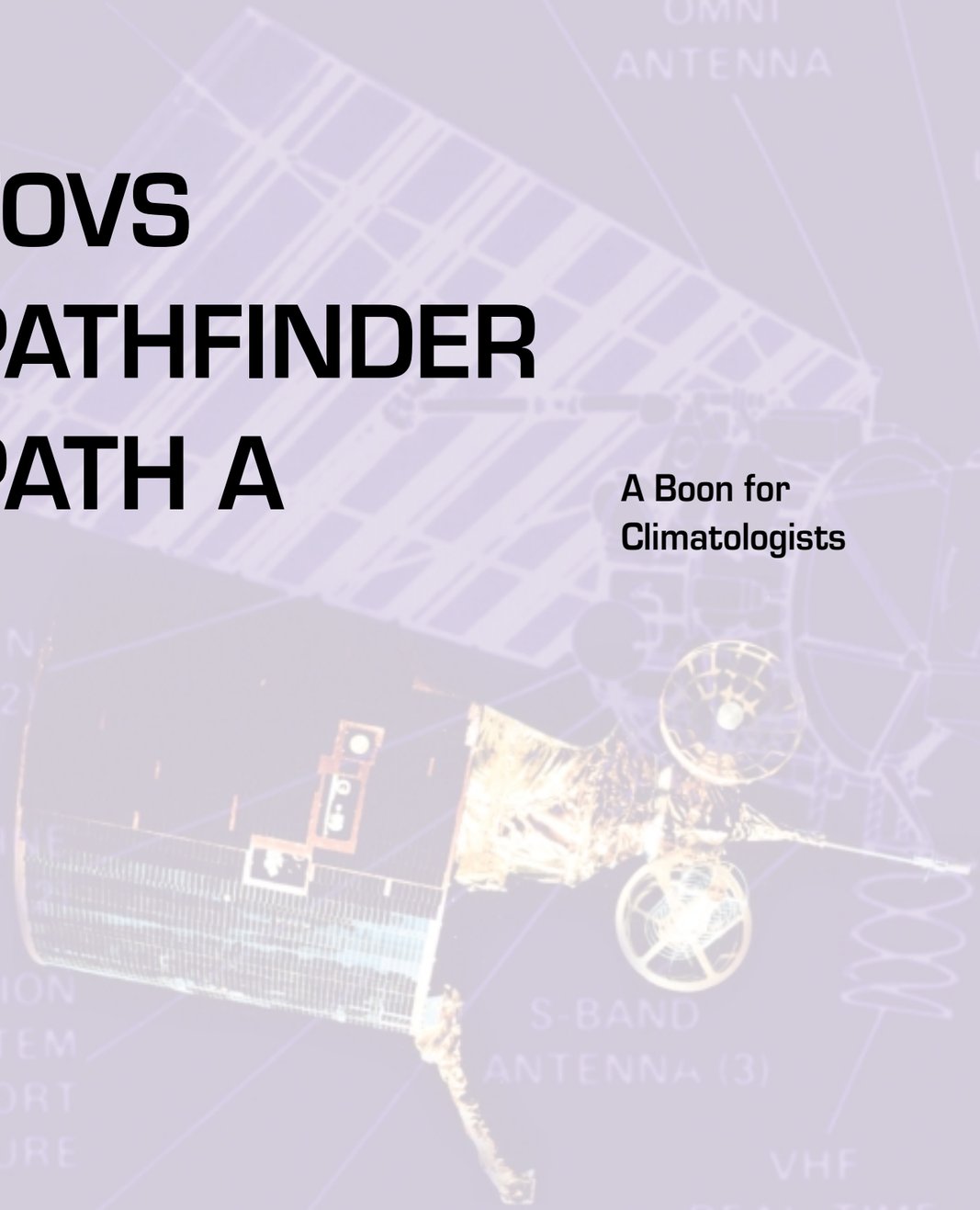
"The Tropics are the heat engine of the atmosphere," says Jagadish Shukla. "That's where most of the energy comes from." The condensation of water is exothermic, which means that when clouds form anywhere in the atmosphere, they release heat. In the big picture, the Tropics are witness to the most massive condensation of water into clouds on Earth. Tropical precipitation releases 75 percent of the energy that the atmosphere uses, and that tropical heat input is one of the major driving forces of air circulation over the Tropics, which drives weather conditions all over the world.

Since Shukla is interested in changes in tropical conditions that affect North America a season or more later, he uses tropical SST data in his models. SST affects cloud formation and directly influences condensation and precipitation in the Tropics. SST changes more slowly than conditions in the atmosphere, and can affect atmospheric circulation and weather over North America many months later. El Niño is the most famous example of this process. The discovery that the warming of tropical Pacific SST serves as a warning of aberrant weather demonstrates the predictive power of SST.

Photographs courtesy of NOAA.

Two sources of SST data—ocean buoys and ships.

TOVS PATHFINDER PATH A



GOES and TIROS satellites, courtesy of NOAA.

A Boon for Climatologists

Is it cooler than normal or warmer? Are we having an El Niño or a La Niña? How intense is it, and how is it affecting the atmosphere?

These are the sorts of questions that climate studies hope to answer. Reaching further into the past than the short memory of the atmosphere and projecting far into the future, climate studies examine trends and changes that take place over decades. Looking at this length of time is necessary to monitor and understand climate variability and to determine if significant trends exist—global warming, seasons of increased flooding, a coming drought.

For climatologists, these studies must have good data sets. Ideally, data would be collected continuously over a period of decades without any major changes to the instrumentation used for collection, which would introduce incongruities and make trends hard to follow. Also, the data should be in the form of time- and space-averaged measurements or estimates convenient for climate studies.

One researcher at GSFC has created such a data set using the facilities at the NCCS. TOVS Pathfinder Path A is a 21-year data set based on measurements taken from several National Oceanic and Atmospheric Administration (NOAA) satellites flown from 1979 through 1999.

Since 1979, NOAA has outfitted its modern generation of low Earth-orbiting meteorological satellites with an instrument package known as the TIROS Operational Vertical Sounder

(TOVS). The TOVS package includes three atmospheric sounding instruments: the High-Resolution Infrared Sounder unit (HIRS2), the Microwave Sounding Unit (MSU), and the Stratospheric Sounding Unit (SSU). These instrument units are all radiometers, or detectors, that are aimed from the orbiting satellites at "spots" on Earth below the satellites' paths. The spectrometers collect all the low-energy photons within a set of spectral intervals or bands rising from the spot below to the TOVS satellite above. Each spot measurement is also a sounding of the atmosphere, counting low-energy photons coming from various layers of the atmosphere.

The photon counts can be retrieved, or interpreted, by an algorithm that translates the counts into geophysical observations—such as the "skin" temperatures of the ground or ocean surfaces, atmospheric temperature profiles, atmospheric water vapor profiles, outgoing longwave radiation, and total atmospheric ozone. A further manipulation of these counts allows the cloud fractions and cloud-top pressures to be calculated.

For the past several years, the Sounder Research Team of Goddard's Laboratory for Atmospheres has used the resources of the NCCS to create a long-term data set of these retrieved geophysical observations from NOAA TOVS satellites for the U.S. scientific community.

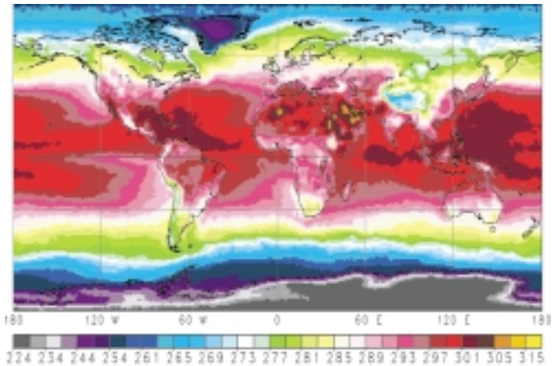
This data set, called Pathfinder Path A, contains the retrieved geophysical observations from the photon counts for the combined HIRS2 and MSU measurements taken over recent decades and represents a detailed structure of the atmosphere integrated over height and averaged over time.

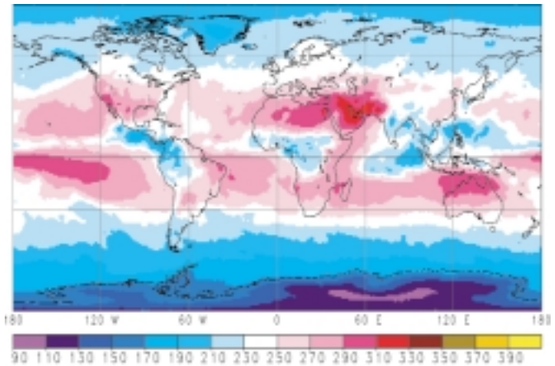
Construction of the data set starts with calibrated level-1 raw data from NOAA, which have already been checked for quality. The raw data are individual soundings and spot-by-spot photon counts from which a gridded Pathfinder Path A data set is produced. The Pathfinder Path A data set averages the soundings on a 1° latitude by 1° longitude grid over 66 levels of the atmosphere to produce a space-averaged and time-averaged data set in daily, 5-day, monthly, and yearly means. The process whereby the raw data are retrieved into the Pathfinder Path A data set first involves comparing calibrated raw data values to values computed from GSFC's Data Assimilation Office GEOS–1 forecast-assimilation system (see "Retrieving the Data").

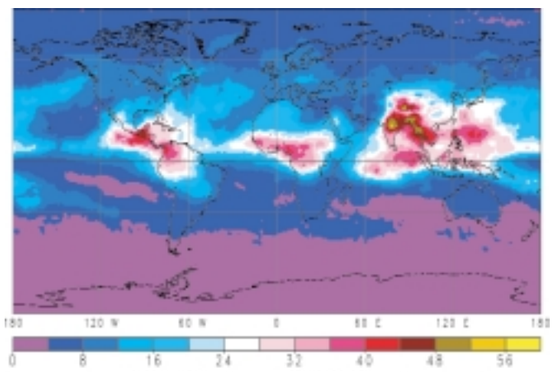
The advantage of Pathfinder Path A is that it uses satellites with extensive global coverage, and, thus, should be useful for diagnosing what the entire atmosphere is doing. "*Doing* is the operative word here," explains Joel Susskind, Senior Scientist of the Sounder Research Team, which has the general charge of developing techniques to infer surface and atmospheric parameters from meteorological satellite observations for use in weather and climate studies. "It's not so much the temperature or any of these other things that we are worried about so much as the variability."

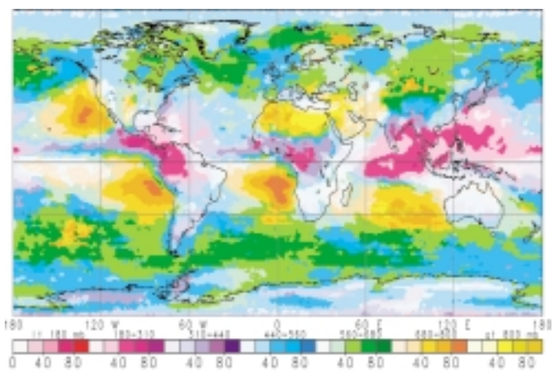
Variability in the atmosphere over short periods of time amounts to changes in the weather. Weather forecasts are just predictions of what the atmosphere will do a few hours or a few days in advance. Over a few months to a few years, atmospheric variability can be observed in the periodic anomalous warming of

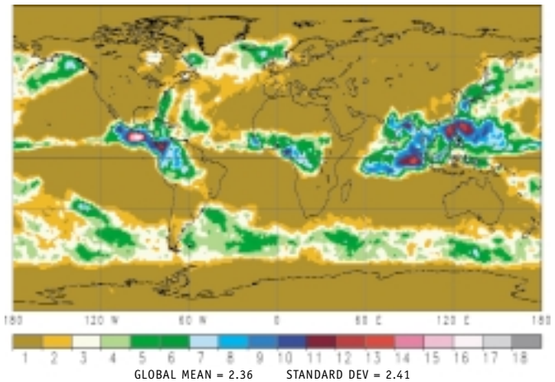
Four views of September 1999. GLA TOVS surface skin temperature (K) (upper left); TOVS-derived outgoing longwave radiation (W/M^2) (upper right); precipitable water (mm^4 10) (lower left); and TOVS effective cloud fraction and pressure (lower right).

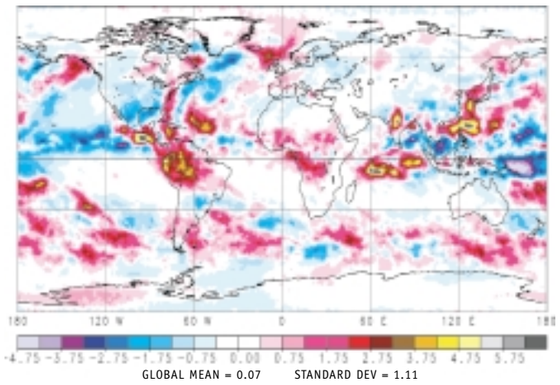












TOVS average daily precipitation in September 1999 (top) compared to TOVS average daily precipitation anomalies, minus climatology (bottom), measured in mm/day.

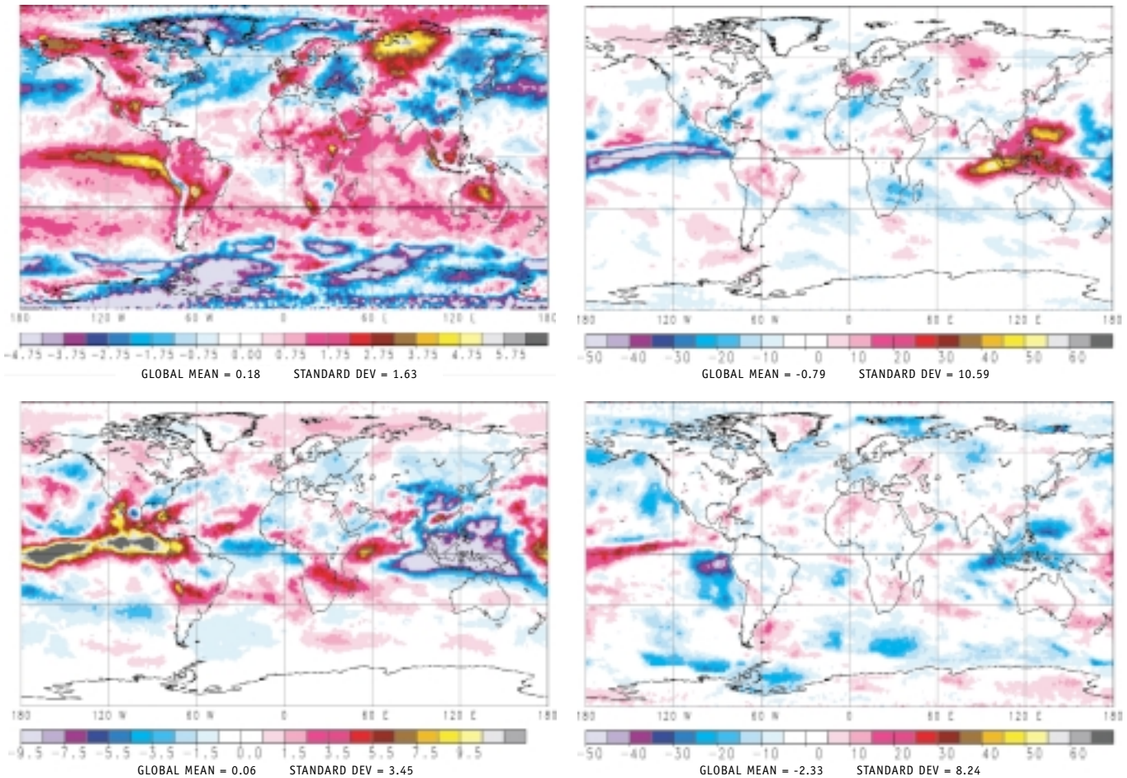
The Pathfinder Projects are a joint effort of NASA and NOAA aimed at producing multiyear climate data sets using measurements from the TOVS instruments.

RETRIEVING THE DATA

The TOVS Pathfinder Path A data set uses as a first guess the algorithm that produces atmospheric temperature and moisture profiles from version 1 of the Goddard Earth Observing System (GEOS–1) General Circulation Model (GCM).

GEOS–1 uses atmospheric soundings and weather data from a variety of sources and produces time- and space-averaged estimates for temperature, humidity, pressure, wind speed, and wind direction. GEOS–1 is a data set of these geophysical properties time-averaged every 6 hours from April 1980 to November 1995. Operationally, the Data Assimilation Office assigns these properties to a space-averaged grid that is 2° latitude by 2.5° longitude with 20 vertical levels.

The Sounder Research Team uses a cruder 4° by 5° grid for constructing the Pathfinder Path A data set. These data are used as a first guess of what the conditions in the atmosphere are like at the time and in the vicinity of any particular TOVS measurement. The retrieval algorithm compares the calibrated photon count from the satellite to values computed from the first-guess value taken from GEOS–1 and employs relaxation routines to find surface and atmospheric states consistent with the observations. Later, some validation tests may be run comparing the Pathfinder Path A data with climate data generated by independent groups such as the European Center for Medium-range Weather Forecasting.



Four views of September 1997 El Niño anomalies minus climatology. TOVS surface skin temperature (C) (upper left); TOVS derived outgoing longwave radiation (W/M^2) (upper right); precipitable water 500 mb to top (mm^4 10)(lower left); and TOVS effective cloud fraction (lower right).

waters in the eastern Pacific known as the El Niño Southern Oscillation (ENSO) cycle. Over even longer periods, observing variability in the atmosphere might answer questions related to climate changes.

Among the data sets that climatologists can use for longitudinal studies are daily surface measurements taken from weather stations and atmospheric soundings taken from weather balloons. However, weather stations and balloon launch sites are few and far between, resulting in insufficient data sampling. Climate studies based solely on these types of data may not be global. "In general, models tend to be better where the [weather balloons] are," says Susskind.

The TOVS Pathfinder Path A data set is not without problems—satellite data are generally less sensitive than those from weather balloons. But TOVS Pathfinder Path A has high-resolution global coverage, and, as of the end of FY99, the data set has grown to more than 20 years and more than 3 TB.

Susskind is using the data from Pathfinder Path A to study trends over a 20-year period to examine questions related to global warming. "We've been concentrating on this interesting question about warming trends in the atmosphere," he says. (See "Applying the Pathfinder Path A Data Set.") The process of generating and maintaining the Pathfinder Path A data set is greatly helped by NCCS's computers charney and soumi. "We used to run at [a rate of] 2 days of analysis per day," says Susskind. "Now we're up to 10 days per day." Each day of climate data added to the data set consumes about a third of a CPU hour.

When the TOVS Pathfinder Path A data set was first created, it covered a 5-year period of MSU and HIRS2 retrievals from 1985 to 1989 collected by the NOAA–9,

APPLYING THE PATHFINDER PATH A DATA SET

One of the issues that Susskind is addressing with the Pathfinder Path A data set is the question of global warming. "To what extent—if any—is there global warming over the past 20 years?" he asks.

There is a contradiction between what surface measurements are showing and what other measurements are showing. Long-term studies of temperature trends at Earth's surface show a warming tendency in recent years. Yet MSU measurements, which are sampled over a 9-km thick slice of the atmosphere, show a smaller warming trend. MSU and surface measurements are not measuring the same thing, however, and the averaging of the MSU measurements accounts for the absence of noticeable warming. Breaking the MSU measurements into smaller steps can bring out the signal.

"Our values are quite consistent," says Susskind. "There is a small warming." Susskind found, though, that this warming is not the same at each elevation and falls off quickly with vertical height.

There are many other uses for the Pathfinder Path A data set. Colleagues of Susskind are looking at the cloud convective index, which is how high and how concentrated the clouds are. By studying the daily averages of the cloud convective index, these scientists hope to better understand the atmospheric properties of air and moisture convection.

Susskind himself also studies Outgoing Longwave Radiation (OLR) to diagnose the atmospheric circulation. OLR is the flux of heat into space and responds to surface and atmospheric temperature and water vapor, as well as the height and amount of clouds. Because OLR changes significantly with the amount of high clouds, scientists use it to estimate precipitation.

Courtesy of NOAA.

The use of satellites to monitor the weather began in a Florida swamp with the launch of TIROS–1 on April 1, 1960.

NOAA–10, and NOAA–11 satellites. Through the 1990s, the Sounder Research Team worked on finishing periods after 1989 and before 1985. In FY99, Susskind used NCCS computers to finish the pre-1985 years, bringing the data from the earliest TOVS satellites into the data set—TIROS N, NOAA–6, NOAA–7, and NOAA–8, which were in operation sequentially from 1979 to 1985.

Another part of the effort involved catching up to the present. As NOAA satellites continue to make TOVS measurements, the Pathfinder Path A data set continues to grow. The DAO uses GEOS–1 to analyze only the period from 1980 to 1995. Susskind continues to use that same algorithm in the analysis of data collected after 1995 because changing it in the middle of a data set could introduce a systematic error, and, thus, the results. He uses SST data from NOAA's Climate Analysis Center and applies the GEOS–1 algorithm to make a first guess for TOVS Pathfinder Path A retrievals after 1995.

Susskind calls the updating process retrospective because data are already old as they are added to the data set. "This is not something designed to analyze today's data today," he says. In fact, the algorithm runs with a 45-day lag, which is the time necessary for the Climate Analysis Center to produce its SST data.

Another issue that Susskind must address is the threat of progress in weather observing systems. Changes in the climate may be quite subtle and represent a small, hard-to-detect signal. Hence, the data used in these climate studies must be consistent through the years. New types of observations, higher quality instruments, and any improvements to the data assimilation algorithm—inevitable progress—may improve routine weather predictions, but they do not help long-term climate data sets.

For example, the instrumentation onboard the NOAA weather satellites changed very little from 1979 to May 1998, when NOAA–15 was launched. With NOAA–15, though, came new instrumentation and a discontinuity in the data—one can clearly see when the new instruments came on line. For this reason, Susskind does not use NOAA–15 data in his data set but continues to analyze NOAA–14 data, which employs the old instrumentation.

"On the one hand, you want to make improvements," says Susskind, "but on the other hand, you want to keep things the same."

An Interview With
DR. JOANNE SIMPSON

Her Career, Meteorology, and Computers

NCCS Highlights caught up with Joanne Simpson as she was preparing for a 3-day conference in December 1999 honoring her lifetime achievements and work.

"I started studying meteorology in 1942. Back then, even being able to put a package of sensors on a balloon to go up to measure the temperature and humidity throughout the troposphere, that was new. And very exciting!"

How have computers been integrated into meteorology—especially from your unique perspective?

Dr. Simpson: Wow! I was one of the lucky people who were able to use the earliest computers. Of course, let me make a general statement about meteorology. All interesting problems in meteorology are nonlinear.

In the old days, before we had computers, we had to linearize the equations that we used in order to make an analytical solution to a problem. The very first work that I did after getting a Ph.D. in 1949 was to study the air flow over a heated island, and it was possible to do that because a heated island could be approximated pretty well by the linear analytic model if we assume that the perturbations in the air flow made by the island were small compared to the overall current that was flowing across the island. We had to do that because there just wasn't any other way. But we realized pretty soon that what we got was, in many ways, realistic. We saw these cloud streets going off the island and extending sometimes a long way downwind. If these clouds that were growing and the cloud streets extending downwind actually started putting latent heat of significant amounts into the atmosphere, then, of course, it immediately becomes a nonlinear problem, when you can't solve it by analytic methods anymore.

Did you use numerical solutions?

Dr. Simpson: Yes, that's exactly what we did, after getting a lot of observations, which in those days was very tedious because it was just pens marking on charts. What we did, believe it or not, was first to make a simplified model of just the vortex, a cloudy vortex ascending, and I managed to reduce the problem to one dimension… and I could solve a simplified version of that on an enormous slide rule that I had. Back in those days, people who got funding from the Office of Naval Research, of whom I was one, were able to obtain slide rules that were about 3 or 4 feet long, and therefore you could do fairly accurate numerical integration, and what I did was

Photograph of tropical cumulus clouds, courtesy of Joanne Simpson. Photograph of Joanne Simpson, courtesy of Deborah L. McCallum, NASA.

numerical integration with height as the cloud or buoyant vortex rose. These results showed a not too bad relationship to observations insofar as being able to predict the height of the cloud and a few simple things like that. But, obviously, you must consider the cloud particles: If water vapor condenses into liquid, the cloud releases latent heat. If the cloud particles grow to become raindrops, they start to fall out of the cloud. You can't go very far with a numerical model without treating equations for cloud particles accumulating into raindrops and falling out.

So in 1954 and 1955, I had a Guggenheim fellowship to work…on cloud study problems, and I went to Imperial College in London, because several of the leading cloud research persons were on the faculty. We started out thinking that the one-dimensional simulation of cumulus clouds is very limited, and we needed to understand how the cloud interacts with its environment. Before we started this, we had been using a very simple way of calculating on a graph how much air from outside got into the cloud. But this time, we decided we should solve the equations on a grid and integrate it. I had a huge piece of graph paper on my desk that had the grid, and we were doing one of the standard numerical integration techniques where you do one point, and that would change, and then that would change all the other points in the neighborhood, and so on.

I had gotten through about the first 1 minute of this job and I had practically erased holes in the graph paper. My former professor, Dr. Carl Rossby, who is probably the greatest meteorologist who ever lived, came to Imperial College on a visit from Sweden, where he had returned from the University of Chicago, where I was his student. He had managed by his tremendous fame and tremendous pull to get one of the earliest computers. This was during the time that ENIAC and MANIAC were just being started at Princeton. Large-scale meteorological problems were for the first time being run in the early 1950s on ENIAC and MANIAC. Rossby had a computer that was very similar.

"Back in those days, people who got funding from the Office of Naval Research, of which I was one, were able to obtain slide rules that were about 3 or 4 feet long."

So Rossby comes in and takes a look at what I am doing and says, "Good God, you won't finish that in your lifetime! You'd better come over and visit me in Stockholm and put it on my new machine." The machine was named BESK, which is an acronym in Swedish.

In the old days, of course, there was no such thing as FORTRAN or any of the languages that you could use with relative simplicity with a machine; you had to program in machine language for that particular machine. I knew that with the time I had left on the fellowship, there was no way I could learn to program in machine language and then carry out all the experiments. But then Rossby said, "Well, don't worry about it. I have several bright young guys at my place, and I'm sure that at least one of them would be very happy to work with you on this problem."

This was a great satisfaction for me, because when I was getting my Ph.D. at the University of Chicago, Rossby said, "Well it's wonderful that a little girl like you is working on cumulus clouds because nobody else is interested in them. They aren't important. And this way you'll be able to stand out."

So I went over there. We started out with a bet between me and Rossby. We had been using the idea of a buoyant vortex to simulate a cumulus cloud. Rossby said that he bet that when I really did this properly, I'd find that it wasn't that easy to get that kind of a vortex, and when I did the problem on his computer we'd see what happened. So for the first few experiments, we had a slightly warm region near the ground, and we followed it to see what happened. We realized that what we should do is not to put the initial slight heat source at the bottom next to the ground, but above the mixed or homogeneous layer we have in the first few hundred meters of the atmosphere and put the perturbation there. Once we did this, we very nicely got a buoyant vortex that was rising naturally.

"Even though it took up the better part of a building, the BESK wasn't fancy enough to be able to put in water drops. So it was a dry model, and it made a vortex very nicely. The only problem was that…about 9 minutes was all we could run the model for before it sort of turned into spaghetti.

"In the old days, of course, there was no such thing as FORTRAN..."

We didn't quite finish it, so a year later (I was working over at Woods Hole at the time), I went back…. One reason that I saw that it was terribly important to finish this was because Rossby was turning 60 years old at the end of that year, and everybody wanted to contribute their best work to a birthday volume for him…. We did finish it and started writing it up, but unfortunately, 2 months after we were there, Rossby died, which was a catastrophe because he was really the last Leonardo DaVinci in meteorology. He knew all aspects of meteorology and was at the leading edge of virtually all the problems that concerned the atmosphere. So…that paper wound up…in the memorial volume to Rossby, which was published on his 60th birthday.

What were the problems, the big questions, when you graduated from the University of Chicago in 1949, in those early decades of your work?

Dr. Simpson: Well, one of the biggest problems was why there are so few hurricanes, and Professor Riehl, who was my thesis advisor, and I worked many years beyond when I got my Ph.D. on how cumulus clouds interact with the large-scale environment and how they act as the main cylinders of the atmospheric heat engine. There was at that time no way to make a numerical model of groups of clouds. In fact, it's only been within the last 3 or 4 years that I've been here at Goddard that it's been possible to make models of groups of clouds interacting with each other, getting all the terms right, and you still have to simplify some things, or you'll even use up the time and space on present-day computers. This modeling effort is led by my colleague, Dr. Wei–Kuo Tao, and he and I have been working together on it ever since he came here, about 12 years ago. He had the beginnings of a cloud ensemble model when he came here, and we have been developing it and finding various ways to initialize the model that are realistic in terms of the forcing of cloud groups and how the clouds interact with each other to make lines, or to make squall lines sometimes, and how they line up relative to the wind.

There have been some beginnings made on the hurricane problem, but hurricanes have so many scale interactions, from almost global down to the size of a cloud itself; it's a tremendously difficult modeling problem, and one that, right now, we don't feel like putting too much ammunition on, because to do nesting from the global model down to the size of a hurricane is beyond what we're prepared to do in our small group here at Goddard. We're not model manufacturers, we're model users, trying to learn something about the atmosphere.

We're sort of waiting for people to finish the scales of models that will resolve clouds or cloud groups on the large-scale models. They started out by only being able to use very coarse grids, like 5° latitude by 5° longitude. It's possible now to do large-scale models with a resolution of down to about 50 km, but it's not fine enough for clouds. For clouds, we have to do at least resolution down to 1 km, and the kinds of cloud models we're working on are called cloud-resolving models, because we actually put enough points in the cloud so that we're mainly studying the process of the clouds, and then we're introducing the larger scale conditions as initial and boundary conditions. I think we're just about finished recoding our models for massively parallel processors—that's the only way. And even so, we're not going to be able to put all of the cloud processes in, because there are processes of forming drops, forming ice, and there are several kinds of ice: There's hail, there's snow, there are tiny little particles, there are great big blobs, and so on. The densities of these are not known, and crystal habits—whether they're hexagons or columns or whatever—what controls them is somewhat known, but it would be very difficult to put into a model.

"It's only been within the last 3 or 4 years that I've been here at Goddard that it's been possible to make models of groups of clouds interacting with each other..."

In connection with the Tropical Rainfall Measuring Mission, we were able to do four aircraft field experiments in 1999, which you have to do in order to try to understand how these particles behave. For the first time, meteorologists now have available a lot of cloud microphysics data, even up to very cold temperatures. These observations are very hard to obtain as they require very expensive instruments and lots of airplane flights. So the modeling in the computer that we've been doing is interactive with observations.

At one time, I drew a diagram to try to explain this: We had a field campaign—this allows us to either develop a theory or to correct a chain of hypotheses that we have—then we work on the computer model and improve it as much as we can, and then we go out in the field with it, or—it's not "we" anymore, actually. It used to be we who would go out in the field with our own airplane back in the 1950s, but nowadays field programs have become so expensive in terms of modern aircraft, digital recording, and modern instruments with remote sensors on most of them. It requires the cooperation of at least several organizations, and very often several nations, to get the kind of data that you need to have. Then you work on that and suggest some things to improve the model and so on. So it's an interaction constantly between the computer models and the observations. I can tell you a very exciting case of that, which we're working on right now.

We worked for 5 or 6 years to get land processes into our cumulus ensemble model. Land processes are very complex because soil moisture varies across the domain, the temperature varies, the evapotranspiration depends on the vegetation, and so on. There's a scientist here in the Mesoscale Modeling Branch who developed a surface processes model, and we worked for years with Dr. Tao and his colleagues to put this together with the cumulus model, because on heated islands off Australia, you can sometimes get clouds up to 18 km that are really spectacular. I've been studying clouds over heated islands since right after my Ph.D. Now we're beginning to get the first runs with really proper surface processes into the model. We're just at a really exciting stage because, in Australia, they've had several field programs to study those huge clouds, so there are aircraft observations and radar observations and wind profile observations. So I think this is one of the most exciting things we've done.

How do centers like the NCCS and supercomputers contribute to the work that's going on at this time?

Dr. Simpson: They contribute enormously! When we want to do a fully three-dimensional run in a large domain, that's the time that we have to use the huge computers.

In some cases, a two-dimensional approximation to a cumulus group or ensemble isn't bad. I started out doing a one-dimensional cloud model, and it just amazes me that in the 50 years I've been working in the field, we've gone from slide rules and horse-and-buggy approaches using World War II old flying boat airplanes that we fastened the instruments onto by hand, to being able to use all kinds of satellite and remote sensors, including Unmanned Aerial Vehicles. I think this is a way that a lot of important observations are going to be made in the future, although piloted aircraft won't ever go out of business. So the development has been absolutely fantastic. I would never have dreamed when I first started my work that we would even be making observations of cumulus clouds from space!

What's very satisfying is that the results are important to people practically, not just because of… forecast computer models, but because of the fact that cumulus cloud groups can make such terrible disasters, like hurricanes, that can cause floods, like the one hitting Bangladesh a number of years ago that killed about half a million people, and also the absence of rain in some situations in El Niño, where regions have just terrible drought. Water is a life-giving substance to almost every creature on the planet, and cloud systems distribute the water around as well as distribute the heat that drives the engine. So it's turned out to be not just a "little girl" thing that only a few people are interested in. By the 1970s, there were whole meetings that were entirely devoted to clouds, and now it's two or three or four different meetings all devoted to clouds, because they've even been broken up into specialties within the study. Clouds are so tremendous in impact to the climate; if the radiative properties of clouds are changed a lot, the greenhouse effect will be changed a lot, so we have to know the radiative properties of clouds to make any estimate of how much of global warming has been contributed by man. It's very clear that there has been global warming, but it's still very controversial whether man has played the major role in that or whether it's something that nature is doing all by itself. So there are vital implications about clouds and being able to measure them.

"When we want to do a fully three-dimensional run in a large domain, that's the time that we have to use the huge computers."

Have there been many changes in the community in terms of how people perceive clouds and in understanding their importance?

Dr. Simpson: Oh, yes! Because of all the types of measurements that are possible, we're able to measure things on a much finer scale than we ever did before—when we first had ENIAC and MANIAC at Princeton, where the great von Neumann was. He was the key person in developing the first meteorological forecast models; they could only do one level, and they could only forecast the flow of that one level. You had to

take the output of that and use some statistical regression methods to say whether it was going to be cloudy or rainy or whatever, and usually the accuracy of those was not very great.

What made you interested in becoming a meteorologist?

Dr. Simpson: Well, I was always interested in the atmosphere and the oceans, and also my father was very interested in aviation. He was the aviation editor of a big newspaper for a long time, and I got to go flying in airplanes with him from about the age of 6. In the first job I ever had, a summer job, I was working for the man who was aviation director of the Aeronautics Commission for the State of Massachusetts, and I started to learn to fly at the age of 16. When I got to the University of Chicago, they had a flying club, and classes that you had to take to get your pilot's license. One was meteorology, and I found it so absolutely fascinating that I wanted to take another class like that. Well, the war was going—World War II—and Rossby at that time was the head of meteorology at the University of Chicago. I wandered into his office to find out if I could take another course, and I came out 5 or 10 minutes later completely committed to being a full-time meteorologist because he said that was by far the best way that I could contribute to the war effort.

> "...there are vital implications about clouds and being able to measure them."

After the war, they were saying, "Now women go home," like Rosie the Riveter. The men were coming back and they deserved the jobs. So when I wanted to go back to graduate school at the end of the war, it was very, very difficult.

But you did!

Dr. Simpson: But I did, and it isn't that difficult any longer, and especially here. One reason I like it so much here at Goddard is that, even by the time I came here, there were enough women scientists who talked science in the ladies room. The critical mass of women is reached when you can talk science in the ladies room! Goddard has always been just great in both science and engineering. It's way ahead of most places.

On May 14, 1997, there was a dedication ceremony in which a Cray T3E was named after you: So how does it feel to have a machine like that churning out 153 billion floating-point operations per second in your honor?

Dr. Simpson: It makes me feel very humble, and also very happy, because usually honors like that don't come to people until after they're dead. There's a machine in here named after Rossby and there's one named after Charney, and they were named after both of them had died, and the fact that while I'm still alive somebody named one after me is an overwhelming honor…. It's very humbling and very exciting that they're having the symposium when I'm alive. When I think of that computer and all the people who are working on it with all their different kinds of problems, it's really very exciting.

Besides having the computer named after you, what would you say has been your crowning achievement?

Dr. Simpson: I got the Rossby Research Medal, the highest research award in my professional society, the American Meteorological Society (AMS). I also was elected the president of the Society, and the AMS is placing me on the cover of the special

issue of the *Journal of Applied Meteorology,* which is dedicated to TRMM. Also, it's unprecedented for NASA to gather people together to have a series of 3 days on a person's work while they're still alive. That's not something I achieved, it's something that other people have recognized. In a way, it's sort of a legend that's gotten larger than life, which makes me feel very humble. Lots of school kids write in and E-mail and ask to write an article about me because I appear in so many of the weather books that they read, so we always send them information and so on.

I mean, one does science by continually failing and trying again, and making the mistakes and doing better. You don't generally just sit there and all of a sudden a light bulb flashes over your head and you say, "Aha! This is the solution to the cumulus cloud problem." It's trial and error, and then learning from observations and learning from the mistakes. I keep trying to tell the young people I interact with, from grade school kids to postdocs, that what you have to learn to be is motivated and persistent and stubborn; it isn't a matter of being a genius.

So speaking of mistakes in science, have you made any?

Dr. Simpson: Oodles! I keep notebooks from the first time I got involved in meteorology. I was asked to contribute my papers to a library that's going to be about pioneering women, so I was going through these notebooks again and explaining them so that people would understand them. Often I'd come to something in the notebook and I'd want to cross the whole thing out because it was wrong. It's not like I made just a few mistakes, either; in science you have to be willing to make them virtually all the time, especially in the old days when we solved things by analytical mathematics. There are just as many but different types of errors one makes in using computers.

I'm still an old-fashioned person, and I like to make as many calculations myself as I can. I like to make actual plots on graph paper because some people have gotten so carried away by machines and computers and computer models that they've sort of forgotten to keep their feet on the ground in terms of observations and plausibility. We have a bit of a conflict in our field now, because most people want to get Ph.D.'s and want to do something with a computer model, because it's easier to change a few things on somebody else's computer model than [to] study another area or another sub-subject or something like that, or than it is to go out in the field on a ship or an airplane or [be] on the ground watching a radar and keeping track of it every minute. If you go out on a field program, say to study cumulus clouds, and there are practically no cumulus clouds during the 3 weeks you're out there (that's happened to me), then you spend all the time and money and don't really have what you want. Whereas, now that there are many kinds of numerical models that relate to [the] atmosphere, to clouds, the temptation on the part of the student is to do those without having had the observational experience that they really need.

> "Goddard...[is] great in both science and engineering.... It's way ahead of most places."

Models are so complex that you can't understand them anymore; an example is hurricane forecasting. My husband has been a hurricane forecaster for years. At the time he started, he and the people working with him still had a tremendous amount

of physical knowledge from having to do these things themselves…. They can at least understand the simplifications that the early models made.

Now there are about four or five models of hurricane motion, and they've been able to improve the forecasts, but still, if there are four models, there are four different tracks, and unless you can use some of the old-fashioned experience, which most young guys don't have anymore, you don't know which one to pick. The models that forecasters use today are so fantastically complicated and there are so many different assumptions in them that you don't know what to do with it. If it doesn't work, you aren't able to say why it doesn't work. The track of a hurricane can now be predicted much better than it was 20 years ago, but the intensity of a hurricane can't.

> "It's not like I just made a few mistakes, either; in science you have to be willing to make them virtually all the time."

Do you have any comments on the future of meteorology?

Dr. Simpson: We can formulate a lot of processes in the cloud models now—for example, radiative processes that we couldn't do before. We can formulate the fluxes of energy from the ocean or the ground into the atmosphere that we couldn't do before, because computers didn't have big enough storage or work fast enough. Every time they get faster, we can do more with the models. On the other hand it's always a conflict between learning how to use the computer and keeping your feet on the ground in observation. And the observations get more sophisticated, too. There are all these optical sensors and remote sensors now that we didn't have, and we have to learn how they work or we can't interpret the results because we don't know what their weaknesses are. As a result, what's happened in my field (I mean meteorology) is, first, we've realized we've had to do environmental science, that the atmosphere doesn't stand by itself, and we've had to look at the ocean and the land and so forth. We also enter in a confrontation with that: We can now put a lot more complex things in the model, but with difficulty. It's very hard to work in two disciplines, to know two disciplines or three. Some fields, particularly meteorology and oceanography, are fortunately so compatible with each other that you can be an oceanographer if you're a meteorologist and vice versa. So we have made great progress in the disciplines between the atmosphere and the oceans, but land processes have been a whole new area for us to try to learn. We can only improve our models by not just treating the atmosphere by itself, but also the environmental factors that are feeding energy into it and taking energy out of it.

So it's a paradox: What we need to know is becoming more and more interdisciplinary, and yet within each subject area everything is more and more complex, so that some people just work in one little area their whole lives.

So to answer your question about forecasting—what's going to be next in meteorology—I just can't say. I hope that we are becoming more interdisciplinary, and I hope that we can do it without becoming superficial. I believe very strongly in the interactions between observations and computer models.

What other developments do you foresee?

Dr. Simpson: Well, I think that there are likely to be breakthroughs in short-term climate forecast; maybe they won't be deterministic forecasts because dear old chaos sets in. It was shown way back in the 1970s that chaos is inherent in the atmosphere. Lorenz showed it. But we can do things like the recognition of when an El Niño pattern is going to occur now, in advance. El Niño[s] can be predicted to a degree of probability quite a bit ahead, and how these affect the regional parts of the globe is becoming better known. So I think that there's going to be huge progress in short-range climate forecasting. I don't think I'd like to stick out my neck in any other area, but scientific processes involving clouds and cloud radiation and cloud properties—they're going to advance. It's challenging because of the difficult measurements that have to be taken, but there's real progress being made there, too.

> "I think that if we keep the measurements and the theories and the models integrated with each other, we're standing on the edge of being able to do even more wonderful things."

I think that if we keep the measurements and the theories and the models integrated with each other, we're standing on the edge of being able to do even more wonderful things.

MICROGRAVITY

Molecular Dynamics Simulations at the NCCS Probe the Behavior of Liquids in Low Gravity

The life of the very small, whether in something as complicated as a human cell or as simple as a drop of water, is of fundamental scientific interest: By knowing how a tiny amount of material reacts to changes in its environment, scientists may be able to answer questions about how a bulk of material would react to comparable changes.

NASA is in the forefront of computational research into a broad range of basic scientific questions about fluid dynamics and the nature of liquid boundary instability. For example, one important issue for the space program is how drops of water and other materials will behave in the low-gravity environment of space and how the low gravity will affect the transport and containment of these materials. Accurate prediction of this behavior is among the aims of a set of molecular dynamics experiments carried out on the NCCS's Cray supercomputers.

In conventional computational studies of materials, matter is treated as continuous—a macroscopic whole without regard to its molecular parts—and the behavior patterns of the matter in various physical environments are studied using well-established differential equations and mathematical parameters based on physical properties such as compressibility, density, heat capacity, and vapor pressure of the bulk material. But certain questions about the behavior of a fluid on the microscopic level cannot be addressed through these conventional means. That's where molecular dynamics comes in.

Photograph of a spreading surfactant-covered drop, courtesy of Joel Koplik.

Molecular dynamics is a technique that employs computers—often requiring the resources and support of a supercomputer center such as the NCCS—to study the behavior of the part in order to shed light on the properties of the whole (see "How Molecular Dynamics Works"). For example, in a molecular dynamics simulation of a drop of water on a Teflon surface, each separate water molecule within the drop is modeled on the computer. Each molecule is a body moving in three dimensions, colliding with the Teflon surface, attracting and being attracted to partial charges on other molecules, influencing the movement of the other water molecules, and together forming the large-scale structure we see as a drop.

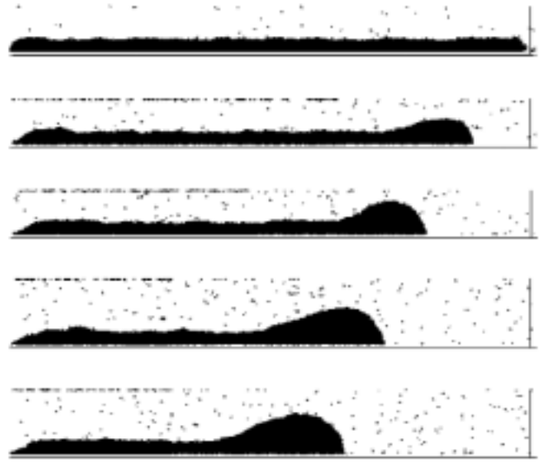
"There are a variety of problems where it is technically relevant to look at a higher scale of resolution where you are not bound by the boundary conditions, and molecular dynamics is the technique for doing that," says Physics Professor Joel Koplik, of the City College of the City University of New York. Koplik and his co-investigator, Professor Jayanth Banavar, of The Pennsylvania State University, are NASA researchers funded through the Microgravity Science Division at Glenn Research Center.

> Without powerful supercomputers like those at the NCCS, such studies would not be possible. The number of calculations required in such a molecular dynamics computation is so large (on the order of 10^{12}) that doing the calculation on a desktop computer or workstation might take several months.

The NCCS supplies the computer resources for their molecular dynamics work, allocating space on both the Cray T3E and a Cray J932se—giving Koplik and Banavar the advantage of being able to study large systems. A workstation might be able to finish a simulation involving about 10,000 molecules in 1 day, whereas the J932se could finish a simulation with 50,000 to 100,000 molecules—and the T3E, hundreds of thousands of molecules—in the same amount of time.

When water is smeared on a waxy surface, it naturally tends to bead into drops; this process is called dewetting. One of the questions Koplik is investigating is the dewetting behavior of thin films of liquids on solid surfaces and, in particular, the rupturing and coalescing of droplets.

In his simulations, Koplik places a film of generic viscous liquid molecules on a solid surface to which they are weakly attracted and observes how the film



Time sequence (from top to bottom) showing a dewetting film at 100, 1,250, 2,500, 3,750, and 5,000 picoseconds.

HOW MOLECULAR DYNAMICS WORKS

In molecular dynamics, a material is modeled as a collection of molecules rather than a smooth continuum with a certain set of properties. Each individual molecule is assigned an appropriate initial position and a random velocity according to a Boltzmann distribution, which means that the average kinetic energy of each molecule is proportional to the bulk temperature, and then allowed to move freely, influencing and being influenced by the surrounding molecules.

An interaction potential energy defines the forces between the molecules, and the computer uses this potential and Newton's equations of motion to calculate the motions of all the molecules. The properties of the bulk of the material—density, temperature, shape, etc.—are revealed through the collective behavior of all of the molecules.

A molecular dynamics simulation first surveys the forces that all of the molecules in the system exert on one another or on the boundaries (like the container), and then integrates these effects forward in time. At each step in the calculation, the computer must move each molecule accordingly for a small amount of time; a small fraction of a picosecond (10^{-12} sec) is standard, because this is the time that it takes for a single atom to oscillate in its microenvironment. Then, the simulation surveys all of the molecules and moves them again and again.

This type of calculation, however, is computationally demanding, which is why the NCCS provides its supercomputer resources. Typical runs may involve at least 10,000 molecules going through thousands of iterations. Each integration involves a minimum of tens of thousands of calculations and may take as long as a few minutes, even on a Cray J932se. In the end, the entire simulation may require hundreds of hours of supercomputer time and produce as much as 100 TB of raw data.

There are many ways to reduce the size of the calculation, but only at the expense of accuracy. The molecules may be modeled as simple geometric shapes, fewer molecules or shorter time scales could be used, or a simplified potential could be defined.

All molecular dynamics simulations have some approximations, because some simplifications have to be made just to make the calculations manageable. Some standard simplifications include assigning particle charges, ignoring certain interactions, or defining a maximum distance for other interactions—such as long-range electrostatic forces. For example, in their investigations, Koplik and Banavar model water as a simplified generic viscous liquid with similar properties and use a perfect crystal to model the solid surface.

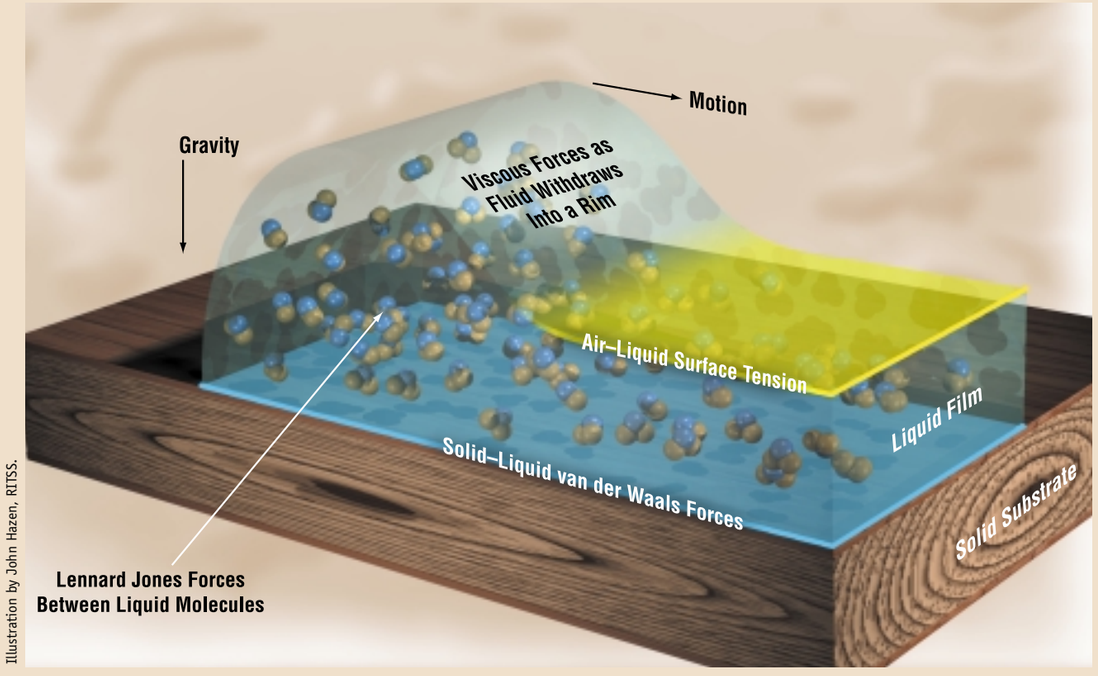
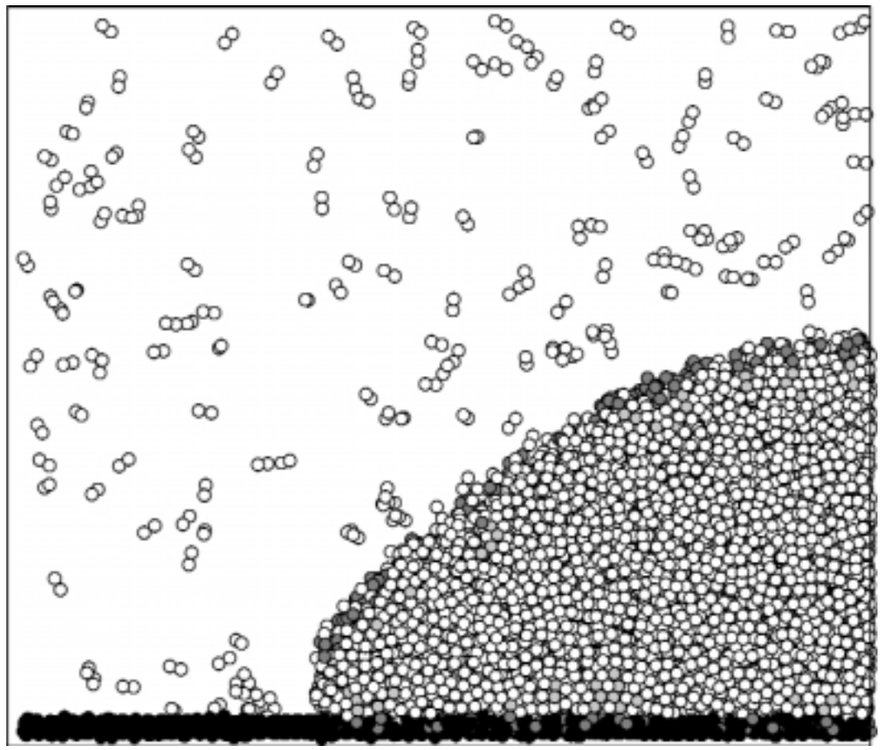


Illustration by John Hazen, RITSS.

Diagram showing the forces that drive a dewetting liquid film on a solid substrate.



A spreading surfactant-covered drop with 10,728 interacting particles.

recedes as the molecules are drawn into a droplet: what the contact area between the solid and liquid is like, what angle is at the interface, and how these things are affected by gravity and the thickness of the film. Rather than laboring to model the water and Teflon molecules exactly as they are, Koplik approximates their size, shape, charges, and morphology to reduce the number and simplify the variety of interactions to calculate.

He also looks at the effect that surfactants—such as soap—have on the wetting process, and what this means for fluid transport. "We're trying to understand under which conditions a surfactant will make a liquid spread where it wouldn't ordinarily," says Koplik. "What, for example, characterizes the phase behavior and the displacement of the surfactant molecules?"

Theoretical studies of fluids may seem academic, but they are of crucial interest to space exploration.

Theoretical studies of fluids may seem academic, but they are of crucial interest to space exploration.

"Wetting issues have an enhanced importance in space," says Koplik, "because they are unmasked, in a sense." On the ground, liquid sinks to the bottom of containers because of gravity, but in space, liquid molecules may have favorable attractions to the container and, without the restraint of gravity, may creep up the sides of the container. Conversely, where, in space, liquids coalesce into extremely large droplets, coalescence is discouraged on the ground, because gravity forces the droplets to flatten out beyond their ability to hold their shape.

In liquids, the forces that act on the molecules are well known. These include surface tension at the air–liquid interface, attractive van der Waals and electrostatic forces between the water molecules, viscous forces that come from the movement and shape of the water molecules, and gravity. Still, the exact nature of such instabilities as surfactant spreading or film dewetting may not be obvious. "Is dewetting," asks Koplik, "initiated by dirt and surface defects or by intrinsic thermodynamic instabilities? Are there patterns to dewetting?" Ground-based theory and experiment can address some parts of these questions, but microscopic information is crucial as well.

Because gravity is a variable in these issues, computational approaches are the only way to study these effects on the ground, and molecular dynamics is one of several possible computational approaches. Though fluid mechanics is often studied at the bulk material level using much less computationally expensive statistical dynamics calculations or Monte Carlo techniques, molecular dynamics simulations are the only way to address time-dependent, nonequilibrium problems such as these.

The Microgravity Research Program, part of the NASA Human Exploration and Development of Space (HEDS) Enterprise, funds researchers' NCCS computer time, which allows them to investigate the physical, chemical, and biological effects of the microgravity environment of space.

THE TORQUE OF THE PLANET

NASA Researcher Uses NCCS Computers To Probe Atmosphere–Land–Ocean Coupling

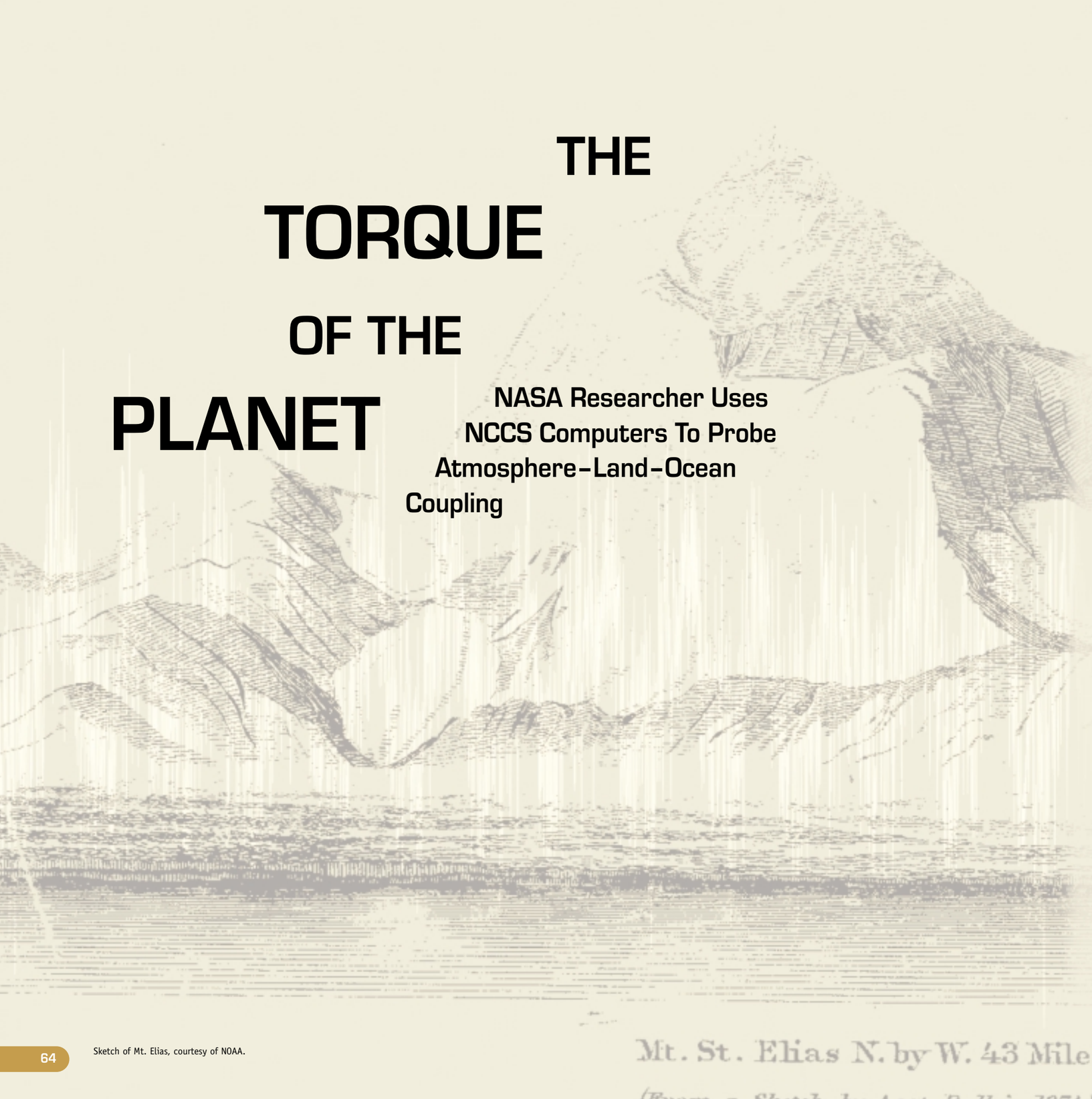
The study of Earth science is like a giant puzzle, says Braulio Sanchez. "The more you know about the individual pieces, the easier it is to fit them together."

A researcher with Goddard's Space Geodesy Branch, Sanchez has been using NCCS supercomputer and mass storage resources to show how the angular momenta of the atmosphere, the oceans, and the solid Earth are dynamically coupled. Sanchez has calculated the magnitude of atmospheric torque on the planet and has determined some of the possible effects that torque has on Earth's rotation.

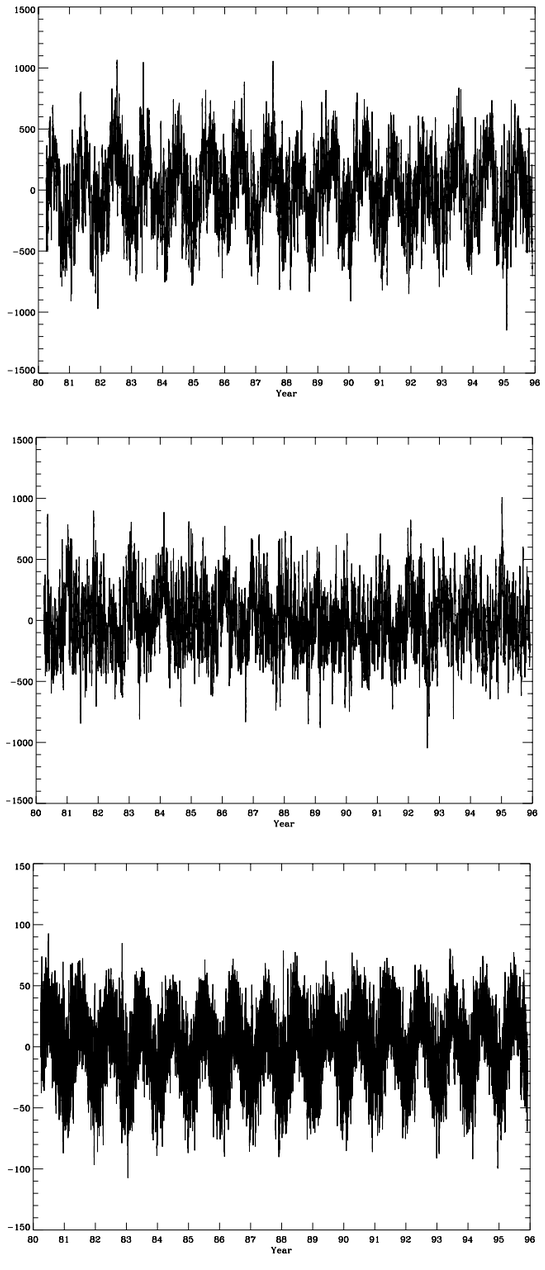
Earth has angular momentum because it has mass and is rotating, which is the same reason that the atmosphere also has angular momentum. Simply put, the air in the atmosphere has mass, the mass of air moves as the atmosphere circulates around the planet, and the circular direction and speed of the mass of the atmosphere yield an angular momentum. According to the law of the conservation of angular momentum, the Earth and the atmosphere must balance each other perfectly.

However, scientists wish to understand more about the nature of the coupling between the momenta of the atmosphere and Earth. Angular momentum accounts only for the mass of the atmosphere or of Earth, without regard for what happens at Earth's surface, where the atmosphere meets the land.

Atmospheric pressure pushes against every vertical object on the land, and that force creates torque. Atmospheric winds drag against the land and sea as they blow, creating torque. These



Sketch of Mt. Elias, courtesy of NOAA.



Time series of total atmospheric torque (pressure plus stress plus gravitational) acting on solid Earth and oceans with mean subtracted and amplitude in Hadleys (10^{18} Newton-meters). X component (top); Y component (middle); Z component (bottom).

Earth has a tendency to undergo slight changes in the speed of its daily rotation because of changes in atmospheric angular momentum. Over the course of the year, the number of hours, minutes, and seconds in a day averages out to be the same, but days are longer than 24 hours in the fall and winter by about 250 microseconds, and they are shorter in the spring and summer by the same amount.

torques then transfer momentum from the atmosphere to the land and oceans, maintaining the momentum balance between Earth's atmosphere and Earth itself. Sanchez has been computing and analyzing these various torque effects using charney, the NCCS Cray J932se.

"We are interested in elucidating how the coupling between the atmosphere and the Earth takes place," says Sanchez. "We want to understand the dynamics of the atmosphere–ocean–solid-Earth system and how torques are transferred from one to another."

The transfer of torque effects can be difficult to understand, because the overall effect of atmospheric torque—the sum of all the various torques over the entire surface of the planet—may be quite subtle. A large positive torque in one hemisphere will cancel a corresponding large negative torque in the other.

Some of the largest torques have no effect on the speed of rotation. For example, the torque that arises from the flattening of Earth at its poles is one of the largest torques of Sanchez's studies. However, because of the geographic symmetry of the flattening, it has no overall effect on the average length of 24-hour days, which vary slightly each season because of other atmospheric torques. Its effect is associated with changes in the orientation of Earth's rotation axis.

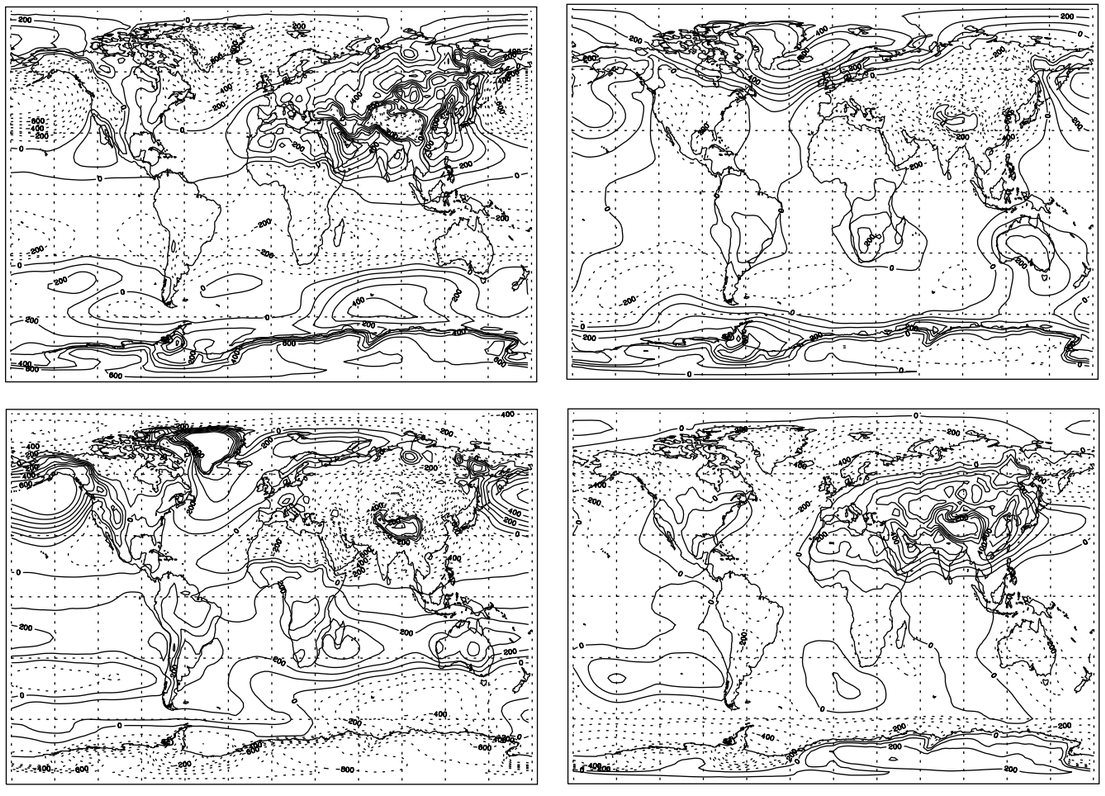
In order to determine how these changes are produced, Sanchez studies 15 years of atmospheric data from GEOS–1, the Data Assimilation Office's first Goddard Earth Observing System general circulation model, which covers April 1980 to November 1995. GEOS–1 data incorporate boundary conditions such as mountain terrain, ground cover vegetation, and seasonal differences in vegetation (e.g., deciduous trees have less wind drag in the winter). From GEOS–1, Sanchez obtains pressure and wind stress data to use as input in his calculations.

The numbers are interesting because they improve our understanding of Earth's rotation. This knowledge is essential for proper navigation on the seas and for satellite orbit determination, since satellites move with respect to a coordinate system defined on Earth's surface, which must take Earth's rotation into account.

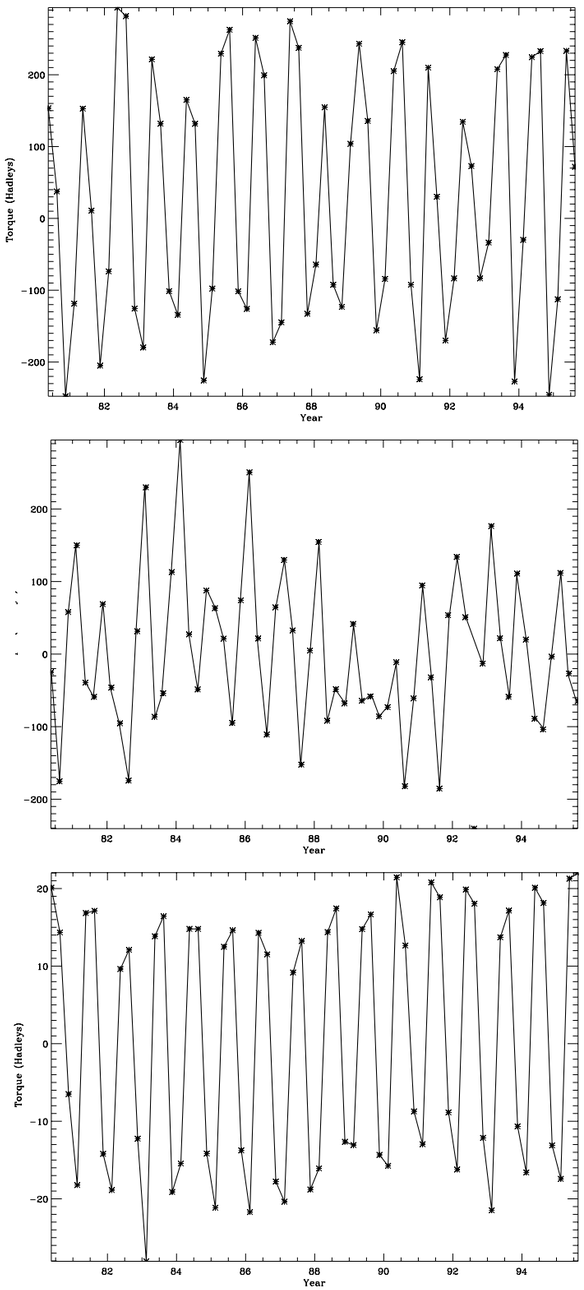
"However," warns Sanchez, "there are things happening that the model cannot define." Because GEOS–1 uses a 2° latitude by 2.5° longitude grid, Sanchez must average topography and vegetation to grid areas that are hundreds of square miles each, approximating land conditions over that entire area, like a low-resolution topographic map. Averaging the landscape smooths the fine details, and, thus, the details of any subgrid torque effects are lost.

"There are ways to approximate those subgrid effects using certain parameters and other models," says Sanchez, "but here they are not present. So you shouldn't set your atomic clock by these numbers."

Timely execution of calculations demands supercomputer time; using charney for his research, Sanchez can model 15 years' worth of data integrated with 3-hour time steps in under an hour. This speed will help in future studies. Sanchez is planning to compare



Seasonal pressure patterns causing the motion of the equatorial torque vector. Pressure differences between four seasonal means minus (clockwise from top left: winter, spring, summer, and fall) annual mean computed for the years 1981–1994. The means are based on 14 averages for each of the four seasons and 14 1-year averages. Solid contours: above-average pressure. Dotted contours: below-average pressure. Pressure measured in N/M^2.



Total atmospheric torque time series of seasonal averages with mean subtracted and amplitude in Hadleys. X component (top); Y component (middle); Z component (bottom).

results from different models that make different approximations. He wants to compare wind stress and pressure data from GEOS–1 and the more recent GEOS–2 and GEOS–3 data sets, as well as look at the existing wave models to see how ocean stress torque is affected by wave dynamics. The Cray's speed should enable Sanchez to compare results across models.

Sanchez will incorporate what he learns into making improved assumptions to use in his calculations. He will also try to increase the resolution of the model so that he can treat the subgrid effects explicitly.

TORQUES OF THE PLANET

When asked to explain what he does, NASA researcher Braulio Sanchez, of GSFC's Space Geodesy Branch, answered, "Oh! This is essentially a calculation of the torque of the atmosphere on the solid Earth and the surface of the oceans."

A torque is a tendency of a force to turn or twist a body about some axis. Atmospheric pressure, for example, can produce torque on a mountain by pushing against it. If Earth were completely barren, a perfect sphere, there would be no torque caused by pressure. If the atmosphere did not circulate, there would be no wind to produce a torque caused by friction.

A torque is a vector cross product, which is a mathematical operation that multiplies together two vectors—quantities with magnitude and directions. The first vector is a radial distance, which is a line pointing from Earth's center to some place on the surface where the second vector is acting. The second vector can be a tangential force, horizontal to the ground, such as wind stress, or a force perpendicular to the topography due to pressure.

As a cross product, a torque is also itself a vector. Physically speaking, this means that torque has some magnitude pointing in a particular direction. Mathematically, cross products are perpendicular to the other two vectors; atmospheric wind stress torque is always parallel to the ground and perpendicular to the direction of the wind.

Atmospheric effects on features of Earth's landscape other than mountains also produce torque. The wind blowing on hills, trees, and even ocean waves constitutes a tangential force and generates a torque.

Not all atmospheric torque is caused by wind friction. Barometric pressure varies over Earth's landscape. Pressure differences on opposite sides of a mountain, for example, will cause the mountain to be pushed in the direction of the lower pressure. The force on the mountain creates a torque along the mountain range. Because pressure differences in the atmosphere can be great, these pressure torques can be quite large as well.

"However," Sanchez notes, "gravitational interactions between the mass of the atmosphere and the solid Earth are sensitive to anomalies in Earth's gravity field, and this dampens the torque due to pressure differences on opposing sides of mountains, canceling about 40 percent of it."

The poles are about 20 kilometers closer to the center of Earth than is the Equator, producing a pressure torque because of pressure differences on opposite sides of the equatorial bulge.

Polar flattening pressure torque has two components, which lie on the plane of the Equator. They combine to produce a torque vector, which lies on the equatorial plane rotating clockwise with the seasons (as seen from the north). This torque vector is responsible for changing the orientation of Earth's axis of rotation.

The resultant of all equatorial torque components is mostly determined by the polar flattening and gravitational torque balance. It rotates clockwise, being in opposition during winter–summer and spring–fall. The maximum magnitude occurs during the spring.

It is of interest to ascertain what seasonal atmospheric pressure patterns are associated with the equatorial torque behavior. In general, the global pressure patterns show an inverse relationship between spring and fall and between summer and winter. In particular, the pressure field over the Asian landmass shows this pattern very dramatically.

SYNTHETIC APERTURE RADAR

The NCCS Enables Search and Rescue

For as long as planes have gone down, dedicated men and women have used ever-improving technologies to aid their search for survivors.

Nearly 2,000 general aviation crashes occur each year in the U.S.—and many, like the Montana incident (see page 80), occur without witnesses. On average, every day in the U.S. one airplane is reported missing. The Air Force Rescue Coordination Center (AFRCC) organizes search missions for about 100 aircraft each year. Some of these are not found before the searches are called off, and are discovered only by chance long after the crash. In some cases, the crash site is never found.

Modern search and rescue operations have become more sophisticated, but still rely on one crucial factor—knowing where to look. NASA's Search and Rescue Mission is using NCCS resources to develop tools for processing radar data that can help these efforts.

United States Rescue Coordination Centers (RCCs), operated under the direction of the U.S. Air Force, Pacific Command (Alaska only), or U.S. Coast Guard, receive and act on distress situations within their area of responsibility.

The AFRCC, located at Langley Air Force Base, Virginia, coordinates overland search and rescue activities in the continental United States. Searches within this region often involve multiple Federal Agencies, such as the Federal Aviation Administration (FAA) and the Department of Defense; State and local authorities; and volunteer search and rescue organizations, such as the Civil Air Patrol.

Georectified radar map of a Montana wilderness, courtesy of NASA/GSFC Search and Rescue Mission Office.

The *National Search and Rescue Plan,* first drafted in 1958, directs search and rescue efforts at the Federal level. Although the wording of the plan has changed over the years, the basic responsibilities remain the same. Signed by the NASA Administrator and the heads of several Federal Agencies, the current plan calls for NASA to apply technology to improve search and rescue. NASA's Search and Rescue Mission Office at GSFC is tasked with this responsibility. The NCCS supports this effort by providing supercomputer resources.

As large a task as coordinating all the people and agencies is, there is often a much larger problem. "The single most difficult task the AFRCC faces when conducting an aircraft search," says Captain Christopher Holmes of the AFRCC, "is determining a search area." Narrowing the search to the smallest area possible is crucial, because the chances of saving victims of aviation crashes decrease rapidly after the first few hours. Any tools that would narrow the search area would be helpful.

"Before the late 1970s, 'eyeball' searching was about the only method available," says Roy Dreibelbis, a consultant for the Search and Rescue Mission at GSFC and a former Deputy Chief of Staff for Operations at Headquarters, Aerospace Rescue and Recovery Service. He has more than 5,000 hours as a rescue helicopter pilot. "When an aircraft turns up missing, we evaluate the weather and terrain, initiate a route search, and hope that a good lead materializes. If the route search is unsuccessful, we expand the search area and hope for a quick find. In many cases, we just don't have a lot to go on."

An important advance in search and rescue in the last few decades has been the NASA-developed Search and Rescue Satellite-Aided Tracking (Sarsat) system, operated by NOAA (see "The Cospas–Sarsat System"). This system, which operates 24 hours per day, 365 days per year, detects and locates transmissions from emergency beacons carried by ships, aircraft, and individuals.

Despite the continued success of emergency locator beacons, they are not without problems. For instance, emergency signals on the 121.5 MHz band are not always reliable, and false alarms plague the computer screens of rescue workers. The Langley AFRCC reports that 90–98 percent of all emergency locator transmissions it receives are false alarms.

An even larger problem is that, in actual plane crashes, beacons may be damaged. "Aircraft beacons are a great idea," says David W. Affens, the Mission Manager for the NASA Search and Rescue Mission Office, "but, unfortunately, in a crash they do not always work."

Visual searches are often delayed for critical hours or days because many aircraft accidents occur in bad weather or as darkness approaches. Even in ideal search conditions, it is still very difficult to locate a downed aircraft, especially in remote, wooded areas where wreckage may be completely hidden by vegetation. Sometimes, broken tree trunks or limbs are the only indication of a crash site.

In the early 1990s, the NASA Search and Rescue Mission Office began investigating the use of Synthetic Aperture Radar for search and rescue (SAR^2). SAR^2 has the potential to complement the Cospas–Sarsat system when a plane's beacon fails to work and to assist visual searches by offering the means to search through vegetation, clouds, bad weather, and darkness (see "SAR^2 Search Scenario" and "How SAR Works for Search and Rescue").

The NCCS is supporting NASA's development of this new technology. "NASA is contributing a lot of things in terms of doing this research and development, and one of them is the Cray computer," says Art Mansfield, Chief Scientist of the NASA Search and Rescue Mission Office. The Mission uses the NCCS's computers and storage facilities because they greatly increase speed of operation and the amount of data that can be processed. "The key here," says Mansfield, "is that without a research facility like the NCCS, we wouldn't have hope to do research like this."

UAVs are a potential platform for a radar-based Search and Rescue system.

When asked to describe the eventual operational SAR^2 system, Affens counted off some possibilities. These included real-time data processing onboard the radar platform, and radar systems on satellites in permanent Earth orbit and on Unmanned Aerial Vehicles (UAVs).

"This system," says Affens, "is intended to be used in remote areas. It's meant to fly over weather, and it doesn't have to be on a piloted aircraft, either. It could be on a UAV." Data could be collected remotely without endangering rescue pilots or pulling personnel away from other aspects of the search.

The NASA group is currently developing advanced algorithms for automatically detecting crash sites, generating maps of those sites, and overlaying Geographic Information System (GIS) data on those maps. A large part of that research and development effort involves the use of the NCCS's supercomputers.

Practical problems including issues relating to radar data collection and storage, and on-board vs. ground data processing are all being examined.

"But," continues Affens, "we are not yet an operational system. We have been called into real searches nine times now, but we are still an R&D system. SAR^2 needs to be tested and applied. We need to put it in the field. We need to produce the SAR^2 evaluation system where real search and rescue people get the opportunity to apply it to real searches."

GSFC's NCCS contributes its sophisticated computer resources to SAR development efforts, which demand a great deal of computer time and space. SAR images require approximately 300–400 floating-point operations per pixel to process, and a typical image may have as many as 22-million pixels.

SAR^2 Search Scenario

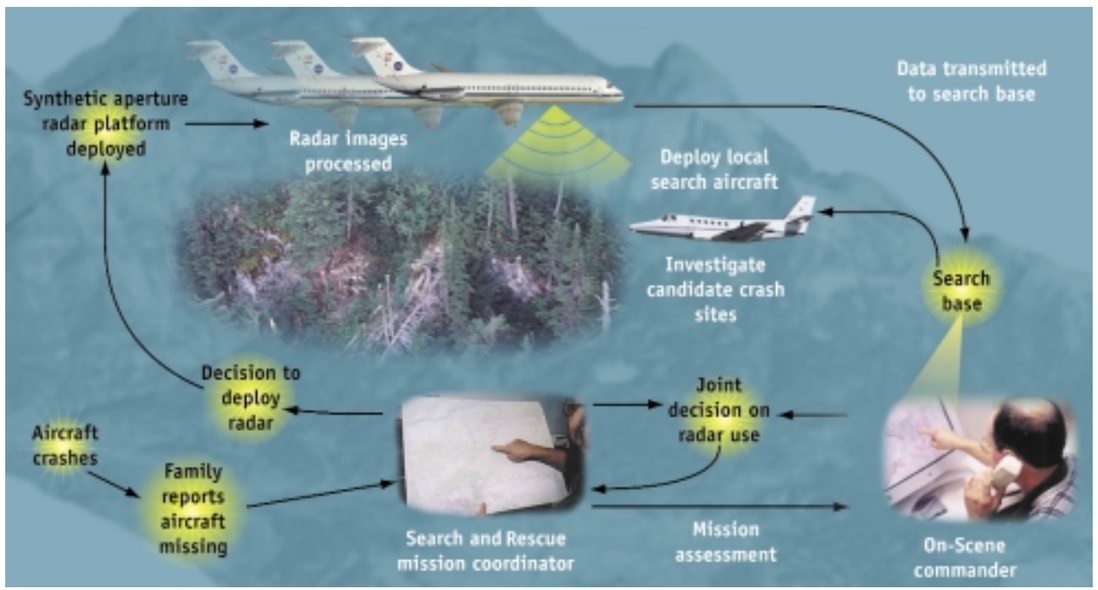


Illustration by John Hazen.

How SAR Works for Search and Rescue

Raw (left), compressed (middle), and autofocused (right) data.

"Synthetic Aperture Radar," according to Affens, "is basically a method that uses computer processing to make radar work better." And when the processing is done on one of the NCCS's Cray J932se supercomputers, it is also faster.

A radar antenna's length determines its aperture, and this, in turn, determines how well individual objects can be resolved. In traditional radars, longer antennas have larger apertures and produce higher resolution imagery. A long, ordinary radar antenna could, in principle, be placed on a search platform and assist in searching for such objects as small planes, but these antennas would have to be impossibly long—several kilometers!

A better solution is to put a smaller radar antenna on a moving body and then mathematically combine separate signals transmitted as the antenna moves, simulating the transmission of radar from a large-aperture stationary source. This technique is called Synthetic Aperture Radar.

SAR is a coherent, microwave imaging system that uses a moving satellite or airplane to transmit radar pulses aimed at the ground. For example, the European Space Agency's (ESA's) ERS–1 satellite has a 10-meter SAR antenna that transmits around 1,700 pulses per second at objects on the ground while it passes 850 km overhead. The processed radar returns from the object have a synthetic aperture length that mimics a 4-km-long stationary antenna.

The motion of the plane or satellite, combined with the wide beam of the radar, covers a swath along the ground, allowing a large area to be searched quickly—typically 256 square km. If a few dozen of these swaths are collected, an area of several thousand square km will be covered. At a search rate of 15,000 square km/hour, many of these areas can be searched in a single afternoon. Typically, a crash search area can be scanned electronically by a SAR in less than a day. The NCCS has been working with Affens's group, providing supercomputer time during the development of SAR for search and rescue to speed up the processing of the images that are collected so that candidate crash sites can be identified quickly.

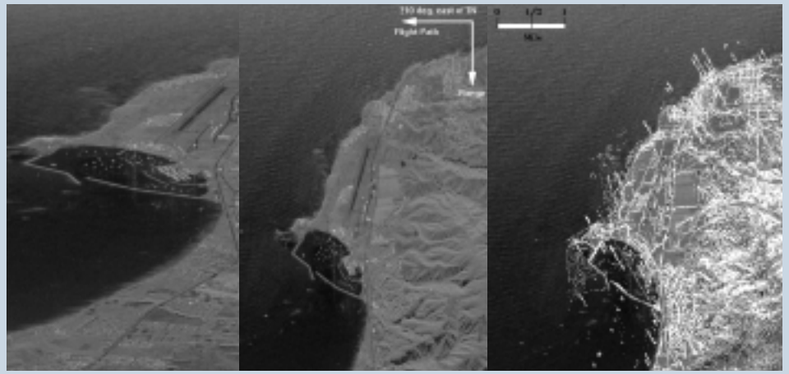
The wide beam of the radar transmission combined with the trajectory of the plane or satellite along its path allow a single object, such as a crashed plane, to be pinged by the radar signal many times as the source passes overhead. For example, the Jet Propulsion Laboratory's (JPL's) AirSAR system can generate more than 12,000 individual returns from each individual ground object at longer radar wavelengths. Returning signals are combined in such a way as to synthesize a single radar return from an extraordinarily sized antenna as long as 4 to 6 km.

SAR devices can operate over a wide range of microwave frequencies, from 1 cm to more than 100 cm, and different frequencies have different imaging properties. Operating SAR at P-band (around 70 cm wavelength or 450 MHz) or at L-band (around 25 cm wavelength or 1,250 MHz), which are relatively long radar wavelengths, allows penetration through trees to the ground. Penetrating ground foliage is essential for search and rescue missions, since planes often crash in forested areas, but this comes at the expense of signal processing complexity.

Returning signals are collected by a coherent radar receiver, which uses electronic processing to preserve the time synchronization of the returns, adding them coherently and creating a sharply focused image.

Then, special range compression and azimuth compression filters turn the raw data into a precise, high-quality image of the ground by exploiting the Doppler signatures, a capability that ordinary radars do not have. The processed image is then focused even further by using a special "auto focusing" algorithm capable of handling longer wavelengths. This technique is a variant of a phase gradient algorithm that employs advanced motion compensation techniques and was first developed by the NASA Search and Rescue Mission Office. The NCCS supercomputers greatly increase the speed of the processing.

SAR is also polarimetric, which means that each radar return actually has four components—HV, HH, VH, and VV—that correspond to the four distinct polarization states of the radar pulses. These states are used in a special process to detect airplane crash sites in the radar image. In addition, through application of a technique called "polarization whitening," the four polarization images are combined into a single, sharper image, which is used to aid the operator in evaluating the result provided by the computer.



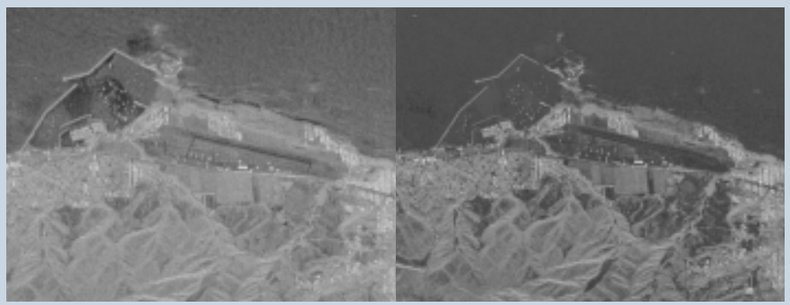
Before georectification (left), after georectification (middle), and georectification with map of area superimposed (right).

The images are large, with each frame typically having around 16,384 x 4,096 pixels for each of the four polarizations as raw data. Each frame may also have more than one frequency band associated with it—typically three—so that a single frame might actually comprise 12 high-resolution images. Each step of the processing produces more frames, and while there is some image reduction along the way, the final frames are still quite large: 14,336 x 1,536 pixels for each of the four polarization states. Each image is saved to the NCCS's UniTree system; then, the particular frames of interest are selected, and the rest are deleted.

The purpose of collecting these radar images is to identify probable crash sites using polarimetric detection. Also developed at NASA, this technique examines the individual pixels of the refined radar image. Each pixel is subjected to a calibrated pattern recognition by the Cray processor to determine whether it could contain parts of a downed aircraft. When a pixel contains a shape characteristic of part of an airplane, the computer flags the location. Certain shapes commonly found on manmade objects, such as metallic perfect right angles, are rare in nature and are, therefore, good indicators of downed planes.

The computer uses the Constant False Alarm Rate (CFAR) technique to maintain only the best returns so that the rescuers will have a manageable number of locations to investigate. The entire automated search process is guided by a search and rescue expert who will be able to eliminate obvious noncrash sites by applying his/her expertise and considering other ancillary information, such as flight path, terrain, and weather.

Another process conforms the final result of the processing to a known geographical map: georectification produces a 3–D map of the actual terrain, providing rescue pilots with a visualization that they can use to navigate to the potential crash sites. The end result of a search might be a few frames, with all of their associated polarization states and frequency bands, that contain one or more possible crash sites.



Before polarization whitening (left); after polarization whitening (right).

The Montana Crash

Aerial view of crash site.

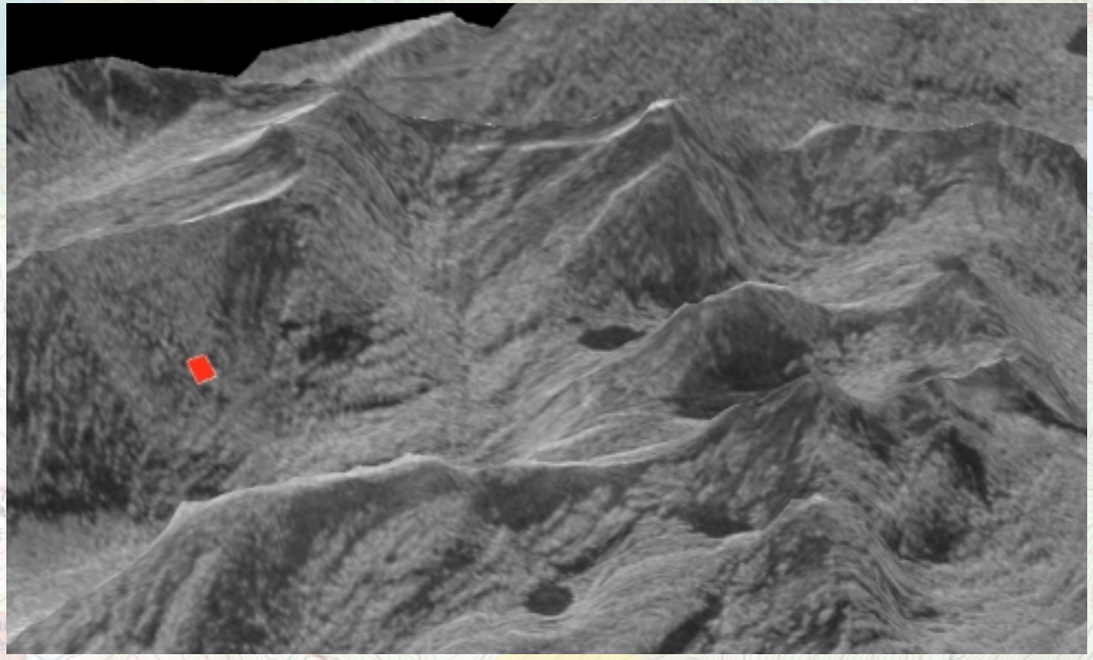
On April 11, 1998, a small Piper Malibu single-propeller plane is flying over a remote stretch of rural backwoods in the mountainous Flathead National Forest, about 10 miles east-northeast of Bigfork, Montana. The plane is at cruising altitude, several thousand feet above the treetops.

The pilot is experienced and in communication with an air traffic controller in the nearby Air Route Traffic Control Center (ARTCC) in Salt Lake City, Utah. He had filed his flight plan the evening before when he and his single passenger left the airport at Madison, Wisconsin, en route to Glacier Park International Airport in Kalispell, Montana. At 2:21 AM on April 12, he radioed that he was beginning his assigned instrument approach to the runway at Kalispell.

And then, silence.

The Piper Malibu belongs to the broad category of general aviation aircraft—small, two-to-six-seat planes flown by a collection of enthusiasts, independent pilots, and small charter operations. They take off and land at remote air strips and major airports all over the country, accounting for some 27 million total flight hours in 1998 alone.

As always happens in the wake of an overland aviation disappearance, the AFRCC initiated a search for the plane soon after it disappeared from radar and failed to land at Kalispell Airport. For days, State of Montana aviation officials and Civil Air Patrol volunteers flew over the areas near where the last radar contact was detected. They searched, but they found neither traces of the missing plane nor clues as to why it disappeared—perhaps an instrument failure, human error, or some unexpected rough weather. The plane simply disappeared—off radar, out of contact, and into the night.



SAR map of crash site.

Flathead National Forest offers a haven for hikers seeking refuge from urban sprawl, with open air, tall trees, towering terrain, and no human contact for miles and miles. But for a rescue mission, it combines the most difficult elements: mountains, trees, and inclement weather through many months of the year. A plane could go down in these woods and never be found, and, for a while, it seemed that the Piper Malibu would become such a statistic.

Then came some unexpected help from NASA in the form of an airborne radar platform from California and one of the NCCS's own supercomputers.

On May 7, the personnel at NASA's Search and Rescue Mission were enlisted to help find the plane. Using AirSAR, JPL's experimental SAR system mounted on a NASA DC–8, they collected radar data over possible areas suggested by the plane's last known heading. Then, using state-of-the-art software developed at GSFC (see "How SAR Works") and one of the NCCS's Cray J932se computers, they processed and analyzed the radar data, generating more than 500 GB of image information. During its months of peak usage, in fact, NASA's Search and Rescue Mission team was one of the largest users of the NCCS resources. The team identified 14 possible crash sites and provided their locations to a private search team hired by the missing pilot's family.

On June 15, the downed Piper was discovered independently by a Montana search pilot near one of the potential crash sites selected by the NASA Search and Rescue Mission Team.

THE COSPAS–SARSAT SYSTEM

Emergency radio beacons, carried by maritime vessels, aircraft, and individuals, are used in situations of "grave and imminent danger" when lives are at risk. Emergency locator transmitters (ELTs), carried by airplanes, were originally intended to be detected by overflying aircraft. Often crashes in remote areas were not found because of infrequent aircraft traffic in these regions. NASA engineers determined that it might be feasible to locate the position of emergency beacons from space and started investigations in late 1960s. The NASA technology was developed into the multinational, humanitarian Cospas–Sarsat Satellite-Aided Search and Rescue system in the 1970s. The first satelliteborne repeater was launched on a Russian satellite in 1982, followed by the first repeater on a TIROS weather satellite in 1983. The first saves occurred almost immediately.

The United States Mission Control Center (MCC) in Suitland, Maryland, operates the U.S. component of the Cospas–Sarsat system. Along with its Russian-managed counterpart, this system has saved more than 11,300 lives since its inception in 1982. Repeaters on satellites relay emergency beacon signals to ground stations, where the approximate location of the beacon is calculated by Doppler shift. This information is then forwarded to the MCC in the country that operates the ground station. The MCC then validates the emergency and forwards the information to the appropriate RCC, which directs the search.

There are now more than 800,000 locator beacons in use worldwide. Improvements in recent years have included the personal locator beacon, which is small and light enough to be carried in a hiker's backpack. Another major improvement has been the development of self-locating beacons. These use the Global Positioning System (GPS) receiver to obtain a precise location and then include the coordinates in the emergency beacon's transmission.

THE EARTH GRAVITATIONAL MODEL 1996

The NCCS—Resource for Development, Resource for the Future

For centuries, men have attempted to understand the climate system through observations obtained from Earth's surface. These observations yielded preliminary understanding of the ocean currents, tides, and prevailing winds using visual observation and simple mechanical tools as their instruments. Today's sensitive, downward-looking radar systems, called altimeters, onboard satellites can measure globally the precise height of the ocean surface. This surface is largely that of the equipotential gravity surface, called the geoid—the level surface to which the oceans would conform if there were no forces acting on them apart from gravity, as well as having a significant 1–2- meter-level signal arising from the motion of the ocean's currents.

The ocean's height varies because dynamic forces are constantly moving the waters; these variations have two timeframes, one that is very long term and, thereby, nearly stationary, and another that is variable. This variable part is dominated by the semidiurnal and diurnal tides, but there are other wind- and thermally driven changes in the ocean current systems that cause ocean surface height changes. Altimeters cannot discriminate directly between the height caused by gravity and that caused by winds, circulation, and other effects: they can measure only their sum. However, this separation is critical for isolating the ocean current changes so important for understanding the climate system.

In 1996, an interagency group of scientists using the supercomputing resources of the NCCS developed the Earth Gravitational Model 1996 (EGM96) to help define the geoid.

Diagram of orbital path of TIROS satellite after launch, courtesy of NOAA.

Since its release to the public, not only has EGM96 been used to improve the model of the geoid, but also has become a useful tool for scientists in many different disciplines, from oceanographers to meteorologists to geophysicists. Says Steven Klosko, a project manager with Raytheon ITSS, "It's been a tremendous aid in understanding ocean and solid-Earth processes." NCCS supercomputers and other resources continue to provide the tools for these researchers.

If there were no tides, and if the waters of the ocean did not circulate, the ocean's surface would conform to an equipotential surface of Earth's gravity field—the geoid. "The geoid is the surface that you would have traced if you were to run a level over an imaginary ocean with no currents and tides," says Nikos Pavlis, a Raytheon ITSS geodesist.

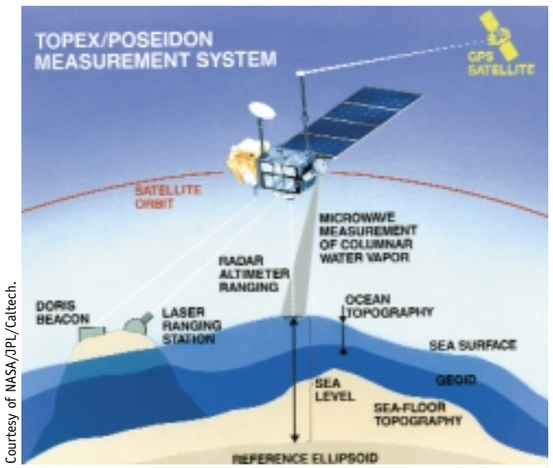
The geoid surface itself has many topological variations, which are related to the density structure of the inner Earth and tectonic features on the ocean floor, such as ridges, trenches, and seamounts. Without currents, tides, and winds, the surface of the ocean would reflect these features exclusively, and, according to geophysicist Erricos Pavlis of the University of Maryland, Baltimore County, and the GSFC Laboratory for Terrestrial Physics, to a first-order approximation, it does. "If you go beyond that approximation, though, there is a departure."

This departure—called dynamic ocean topography—is a manifestation of general ocean circulation and currents. The long-period changes that occur in dynamic ocean topography over decades and longer are one of the keys to studying an evolving climate system, given the importance of ocean circulation in the redistribution of Earth's heat from the equatorial regions. These studies have a central role in climate predictions.

The question then becomes how to separate the dynamic ocean topography from the geoid. A reference ellipsoid is used as the reference surface in these calculations. Geoid height is the distance between the ellipsoid and the geoid, and the sum of the geoid height and the dynamic ocean topography is the height of the ocean at any given point above the reference ellipsoid. Altimeter-bearing satellites such as ERS–1, ERS–2, and TOPEX/Poseidon measure the distances between their orbital position and the surface of the ocean. Accuracy for determining the dynamic height depends on an elaborate orbit determination process using precise tracking of these satellites. Current satellites, such as TOPEX, allow these surface heights to be determined to an accuracy of 2 to 3 cm—as good as most tide gauges for measuring the height of the ocean.

However, without a way of separating the dynamic ocean topography from the geoid, studies of ocean circulation and dynamic ocean topography would be impossible from altimeter data alone. Accurate geoid height determination is critical to the extraction of dynamic ocean topography information from altimeter data.

EGM96 was developed by a group of U.S. Government and contractor scientists working at GSFC in collaboration with the National Imagery and Mapping Agency (NIMA) and The Ohio State University. Using NCCS computing facilities, this interagency group created an improved mathematical model of Earth's gravitational field that can be used to calculate (among other geodetic quantities) the geoid height. Researchers used satellite tracking, altimetry, and direct measurements of the gravity acceleration on land, sea, and air to derive a solution based on these data (see "Earth's Gravitational Field Determination").



Courtesy of NASA/JPL/Caltech.

Diagram of the TOPEX/Poseidon measurement system.

The calculation is computationally demanding—requiring approximately 1 trillion calculations to estimate an orbit for a single day—a task well suited to the NCCS's supercomputers.

EARTH'S GRAVITATIONAL FIELD DETERMINATION

Satellite paths show the variations in Earth's gravitational field. Their orbits change slightly according to the changing attraction of the mass below. With precise knowledge of the gravitational field, of the position and velocity of a satellite in orbit at some point in time, and of all the other nongravitational forces acting on that satellite—solar radiation pressure, atmospheric drag, engine thrust—one could predict a satellite's position days in advance with high reliability.

But is the inverse also true? Could one take satellite tracking information—precisely measured arcs of an Earth-orbiting satellite—and calculate the gravitational force acting upon that satellite? Could one then determine Earth's gravitational field to high precision?

That's exactly what an interagency group led by NASA scientists at Goddard, using NCCS resources, was able to do in 1996.

The basis for the development of a global model (EGM96) of Earth's gravitational potential is satellite tracking data. The orbits of more than 40 satellites are carefully tracked to detect any perturbations in their paths that are introduced by small variations in Earth's gravitational field, caused by the nonuniform distribution of Earth's underlying mass and density. The calculation is computationally demanding—requiring approximately 1 trillion calculations to estimate an orbit for a single day—a task well suited to the NCCS's supercomputers.

"You have some idea of how a measurement is related to the gravity field, and you write down an equation with parameters related to the gravity field," says Erricos Pavlis. "The problem is how to take the pieces that are incomplete and inhomogeneous and put the puzzle together." Unlike any ordinary puzzle, though, the gravity field problem has so many pieces and is of such complexity that supercomputers must be used to solve it.

There are approximations used in orbital solutions that somewhat simplify orbit determination and geophysical parameter improvement, but these problems are highly nonlinear. Solving them requires approximating the mathematical model of the gravitational field with a construct called a truncated Taylor series expansion, from which only the first (derivative) terms are retained. The orbital solution is iterated until it converges. Successive iterations yield improvements to the approximate values of the orbit parameters. When a great number of orbital arcs are combined, it is possible also to estimate simultaneously all of the geophysical parameters as well as improved orbits, including spherical harmonic coefficients that describe the global gravitational field model. The EGM96 solution involved thousands of arcs.

A single arc will contain up to a few hundred arc-specific parameters, and 6,000-7,000 parameters that are common to all the satellite arcs (e.g., spherical harmonic coefficients of the gravity field, tidal parameters, tracking station positions), and these are used to determine a normal matrix for the arc. Each of the 40-plus satellites used in the EGM96 solution contributes some preselected set of arcs, although low-orbiting altimetry satellites and those tracked by GPS have many more observations per arc. For example, a LAGEOS 30-day arc might have some 6,000 observations, whereas a 30-day arc for TOPEX/Poseidon could have as many as 60,000. After each arc is represented by a normal matrix, which is usually more than 100 MB, the NCCS supercomputers are used to add all of the arcs into one giant matrix, using linear algebraic techniques, for each satellite. EGM96 included 13,864,000 observations derived from satellite tracking and direct altimeter measurements of the oceans. This number does not include the observations of gravity from surface gravity surveys (see next section).

However, not all data sources are equal. For example, the tracking data from LAGEOS are much more precise than those from GEOS–3 because the LAGEOS satellite is tracked with the superior laser-ranging technology and is in a much higher, more stable orbit. On the other hand, LAGEOS is far less sensitive to the gravitational field than is GEOS–3, because of its much higher altitude (~6,000 km vs. ~840 km). The satellite data must, therefore, be carefully and appropriately weighted.

Determining these relative weights to combine all of the satellite matrices into a unified solution has been the major contribution of NASA's scientists; it is also the most computationally intensive part of the procedure and is the major component of the resources used at the NCCS.

With unlimited computer resources, Variance Component Analysis (VCA) might be used to test separately the multiple arc parameters within each separate arc normal matrix for possible weighting. But with limited resources, each matrix is weighted as a whole. After all of the data have been processed, an objective estimation process is used to access the weight for each satellite's contribution. These techniques rely on making solutions that exclude and include data from each satellite in turn. The weight estimation is, thereby, an iterative process, requiring hundreds of solutions and demanding tens of thousands of CPU seconds each time.

This approach allows scientists to obtain an optimal solution from the inversion of a large normal matrix based on millions of observations that all contain information about the thousands of parameters that must be estimated. It also simultaneously yields an accurate estimate of the solution's error levels, which are obtained from the resulting solution's statistics.

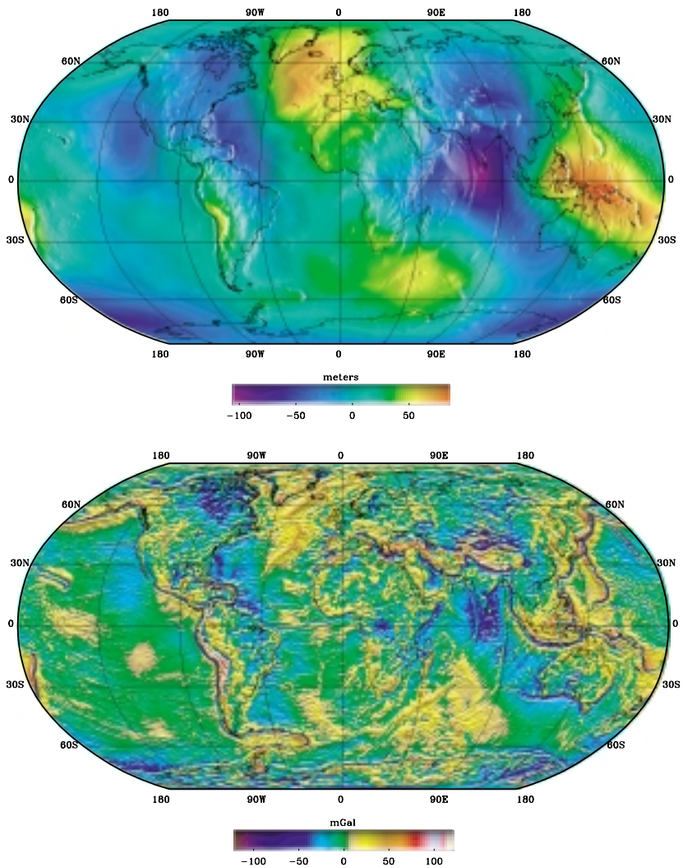
According to geophysicist Frank Lemoine, of the GSFC Laboratory for Terrestrial Physics, "The analysis is based on the pertubations of a spacecraft caused by the densiy variations inside the Earth." These can be derived from the orbit determination process using high-quality tracking data. These variations cause the gravitational potential to deviate from that of a regular ellipsoid, and this is what can be measured on land, over water, or at orbit.

To date, no single source of data can provide both total global coverage and high sensitivity to the fine structure of the gravitational field. So, several sources of data had to be combined to calculate the solution for the global model for Earth's gravitational potential (see "Sources of Data").

EGM96 was made available on the Internet as soon as the model was completed. "Some of our colleagues at the National Geodetic Survey were literally waiting with their hands on their computers," says Erricos Pavlis. Approximately a month after EGM96 was released, the National Geodetic Survey completed a new map of Earth's geoid over the contiguous 48 States.

Today, using NCCS resources, scientists continue to improve EGM96 by incrementally incorporating further satellite tracking data as they become available and improving the quality of available surface gravimetric data. In the next few years, new data will become available from upcoming space missions. These include GRACE (ca. early 2002), which will use two satellites with a continuous microwave link between them to get an accurate measure of gravitational acceleration along their line of sight at altitude; CHAMP (launched on July 15, 2000), which has a GPS receiver and accelerometers onboard and will be tracked by the GPS satellite constellation; and GOCE (ca. 2004), which will use an onboard gradiometer to measure the gravity gradient in three directions. These missions will surpass what is currently possible in high-resolution sampling of the global gravitational field from space, and the NCCS will provide the computational environment needed to translate these new data into greater knowledge of the planet.

Each new data source will be brought into a global geopotential solution, and each will demand more of the NCCS's resources. "The first decade of this century will be the decade of the Earth's gravity field," says Lemoine. "We will be able to map global gravity from space with an unprecedented precision at a higher resolution."



Top Left: The 30' x 30' mean values of geoid undulations computed from EGM96 to degree and order 360. The values refer to the WGS84 (G873) system of constants, which provides a realization of the geometry and the normal gravity potential of a mean-Earth ellipsoid. The permanent tide system is "nontidal," and the units are meters.

Bottom Left: The 30' x 30' mean values of free-air gravity anomalies computed from EGM96 to degree and order 360. Values refer to the WGS84 (G873) system of constants. The permanent tide system is "nontidal," and the units are mGals.

SOURCES OF DATA

"There is not a single measurement that can provide all the answers," says Erricos Pavlis.

Because no single source of data has both enough global coverage and high sensitivity to the gravitational signal, none can be used exclusively to determine Earth's gravitational potential. Three complementary sources of data must be tapped: satellite tracking, satellite altimetry, and surface gravimetry. Each has advantages and disadvantages.

NASA has been tracking satellites in orbit since the early 1960s. Various methods, such as Doppler, laser ranging, radar, Global Positioning System (GPS), and satellite-to-satellite systems (TDRSS), are used today. Some modern satellites (e.g., TOPEX/Poseidon) are tracked so accurately and modeled so well that their locations are known to within a few centimeters at all times.

Ground-based satellite tracking systems (laser and Doppler) cannot track the satellites continuously; satellite-based tracking systems like GPS can. Satellites sense a "smoothed" version of Earth's gravitational field. This smoothing effect increases with altitude; thus, very high orbiting spacecraft are used for determining only the broadest features of the field. The most useful data come from satellites in low orbit, but these rarely are tracked to required precision. For instance, LAGEOS, orbiting at 6,000 km, is far less responsive to the shorter wavelengths in the gravitational field than is GFZ–1, which orbits at 350 km.

Satellite altimeter data have a dual character. They are sensitive to the broad features of the gravitational field, observable through their orbit perturbations, and to the fine details of its structure, sensed through their ocean surface tracing transects. But the incorporation of altimetry data into gravitational field model development is complicated by the fact that dynamic ocean topography must be accounted for or estimated simultaneously. Thus, in EGM96, models representing the broad features of the dynamic ocean topography were estimated simultaneously with the gravitational field parameters.

Gravimetric data are also important in situ measurements defining the EGM96 solution because they help resolve model parameters that would be difficult to separate otherwise. Gravimeters are devices that measure the force of gravity by determining the differences in the weights of a constant mass at different sites. Operational gravimetry is an old technology; data collection is tedious and expensive, because the instruments must be taken to the sites to take the measurements. Hence, many remote areas are almost entirely devoid of observations. Furthermore, political obstacles have made it difficult for the science community at large to gain access to all existing data. However, progress has been great in eliminating this last obstacle.

EGM96 heralded a major improvement over previous models because it was developed with previously classified NIMA data compilations, and many countries made data available to this project for the first time. However, coverage was still sparse over vast stretches of the Southern Ocean, Antarctica, and significant parts of Africa and South America.

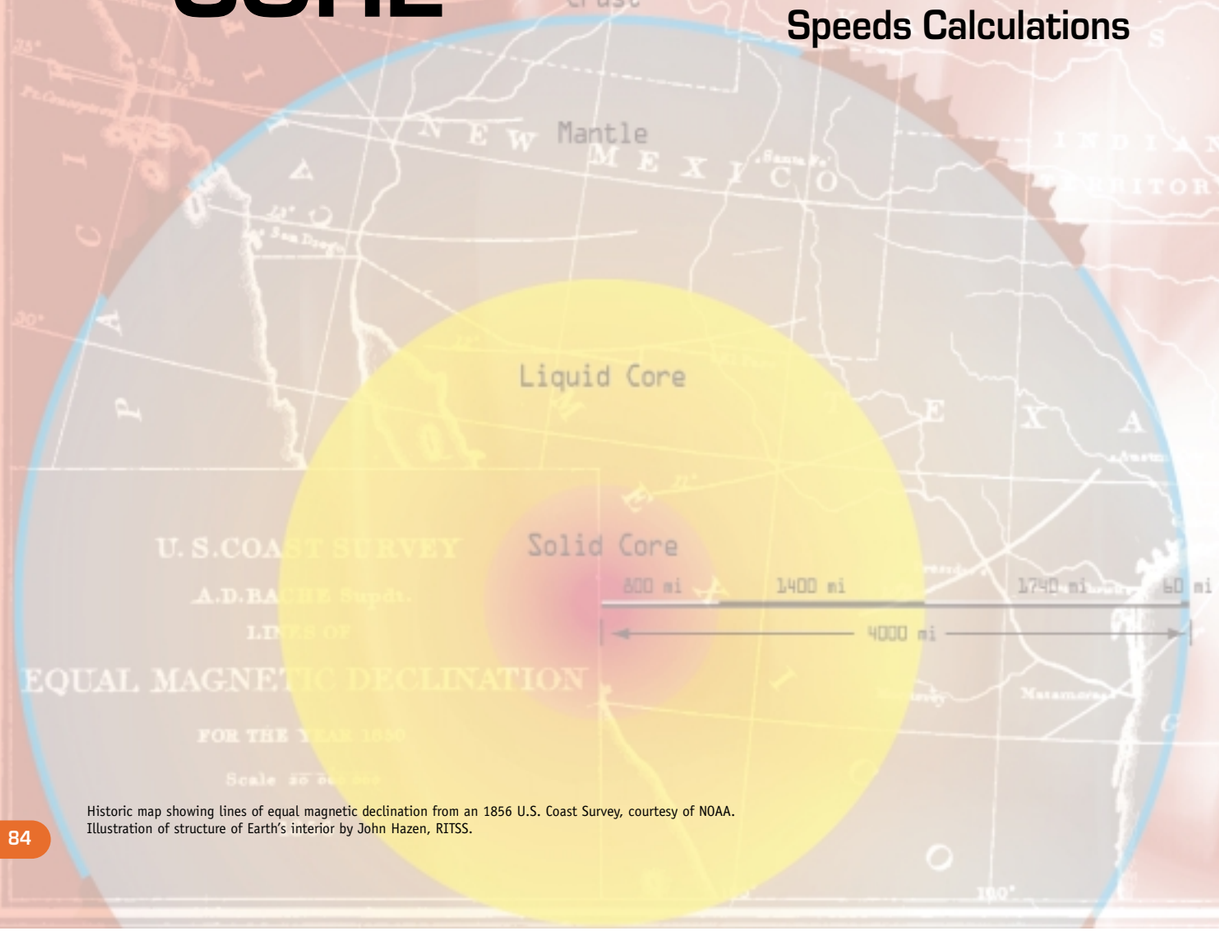
Instrumentation (and possibly preprocessing) differences have made some of the data collected over land suspect, and there are reference frame issues that are difficult to resolve for data coming from bordering countries. Collecting data over water is even more difficult than over land because of errors introduced by the movement of the ships carrying the instruments. Nevertheless, despite all of its shortcomings, surface gravimetry is currently the only source of fine-detail gravitational information over land areas.

"We acknowledge the NASA Center for Computational Sciences (NCCS) at the Goddard Space Flight Center and especially Milt Halem, Nancy Palm, and Tom Schardt for their support. Our altimeter data, and satellite-only normal equations included between 8,000 and 12,000 parameters. The quantity of separate arcs and the size of the normal equations taxed the storage and other operational aspects of the NCCS supercomputers. We are grateful for their patience and support throughout this project, and point out that without their facilities, the EGM96 model could not have been created."

—From *The Development of the Joint NASA GSFC and the NIMA Geopotential Model EGM96*

SIMULATING THE DYNAMICS OF EARTH'S CORE

Using NCCS Supercomputers Speeds Calculations



Historic map showing lines of equal magnetic declination from an 1856 U.S. Coast Survey, courtesy of NOAA.
Illustration of structure of Earth's interior by John Hazen, RITSS.

If one wanted to study Earth's core directly, one would have to drill through about 1,800 miles of solid rock to reach the liquid core—keeping the tunnel from collapsing under pressures that are more than 1 million atmospheres—and then sink an instrument package to the bottom that could operate at 8,000°F with 10,000 tons of force crushing every square inch of its surface. Even then, several of these tunnels would probably be needed to obtain enough data.

Faced with difficult or impossible tasks such as these, scientists use other available sources of information—such as seismology, mineralogy, geomagnetism, geodesy, and, above all, physical principles—to derive a model of the core and study it by running computer simulations. One NASA researcher is doing just that on NCCS computers.

Physicist and applied mathematician Weijia Kuang, of the Space Geodesy Branch, and his collaborators at Goddard have what he calls the "second-ever" working, usable, self-consistent, fully dynamic, three-dimensional geodynamo model (see "The Geodynamo Theory"). Kuang runs his model simulations on the supercomputers at the NCCS. He and Jeremy Bloxham, of Harvard University, developed the original version, written in Fortran 77, in 1996.

The first working dynamo model was created in 1995, by Gary Glatzmaier, now at UCSC, and Paul Roberts, of UCLA, who is Kuang's former colleague and graduate school advisor. Before Glatzmaier, Roberts, and Kuang began their work, scientists

EARTH'S CORE: A CHALLENGE TO STUDY

Modeling Earth's magnetic field was easier 4 centuries ago in Elizabethan England when, in 1600, Her Majesty's Royal Physician, William Gilbert, explained simply that Earth is a large magnet. The notion became a popular one and lasted up to the dawn of the 20th century, by which time several careful scientific observations had called Gilbert's theory into question. His assertion does not account for the variances in the magnetic field—variances in both magnitude and polarity—that occur on timescales from a few years to hundreds of thousands of years. The intensity of the magnetic field varies noticeably over the course of the lifetimes of the scientists measuring it. When the magnetic field intensity was first measured in the 1830s by Carl Friedrich Gauss, it was nearly 9 percent greater than it is today. Also, Earth's magnetic field lines undergo a westward drift of about 0.1° longitude per year, which was recognized as far back as 1698 by Edmund Halley.

In 1906, French physicist Bernard Brunhes noticed an even more dramatic variation in Earth's magnetic field. He studied core samples of volcanic rocks and found that the polarity of the magnetic field shifted over the course of Earth's history, and noticed that rock samples formed in different epochs had acquired opposite polarization.

Gilbert's theory could not explain this phenomenon, either—a permanent magnet cannot reverse its polarity once, let alone periodically throughout geological time. The only possible explanation for polarity reversal would be that the geomagnetic field comes from some process deep within Earth—a dynamic process.

Hence, in 1919, the British physicist Joseph Larmor proposed the geodynamo theory: Earth's magnetic field is generated and maintained by flow in Earth's fluid core. Larmor was the first to believe that the real origins of Earth's magnetic field reside in the core, and scientists have since generally come to accept this theory.

"Many observed phenomena related to the Earth's magnetic field have roots in the core–mantle interactions," says Benjamin Chao.

The problem with the geodynamo theory, though, is that there is no way to study the dynamics of the core directly. Larmor, for instance, had no way of knowing whether his geodynamo theory was correct. It would not be until the end of the century before Kuang and his colleagues had computers—like those at the NCCS—powerful enough to make simulations of the core.

did not know whether dynamo models could actually simulate real Earth processes. Kuang and his colleagues soon realized, however, that some of the approximations they decided to use might be more appropriate to the Earth's core.

Kuang's model begins with as good a picture as scientists have of the core and makes certain assumptions where there are gaps in knowledge of Earth's interior. One cannot study the dynamics of Earth's core by simply observing its most notable effect—Earth's magnetic field—because some parts of the field are completely contained within the core and lower mantle and cannot be seen or be measured by surface instruments (see "Sources of Data"). "The core gives us a very reluctant signal," says geophysicist Benjamin Chao, Head of the Space Geodesy Branch.

To implement his model, Kuang needed to integrate the equations for extremely long periods of time. Since some effects related to Earth's magnetic field occur on timescales on the order of 100,000 years or so—such as the reversal of the magnetic North and South Poles—his model must be simulated over that same timescale, and the output must be able to show such aberrations. However, extremely long simulations are computationally difficult because of the number of time steps involved.

Kuang uses a combination of pseudo-spectral and finite difference methods for numerical integration and runs his model in packets of about 15,000 to 20,000 years at a time, which take an average of 5 days of CPU hours running on 8 processors of one of the NCCS's Cray J932se supercomputers. This is where the NCCS's contributions are so important.

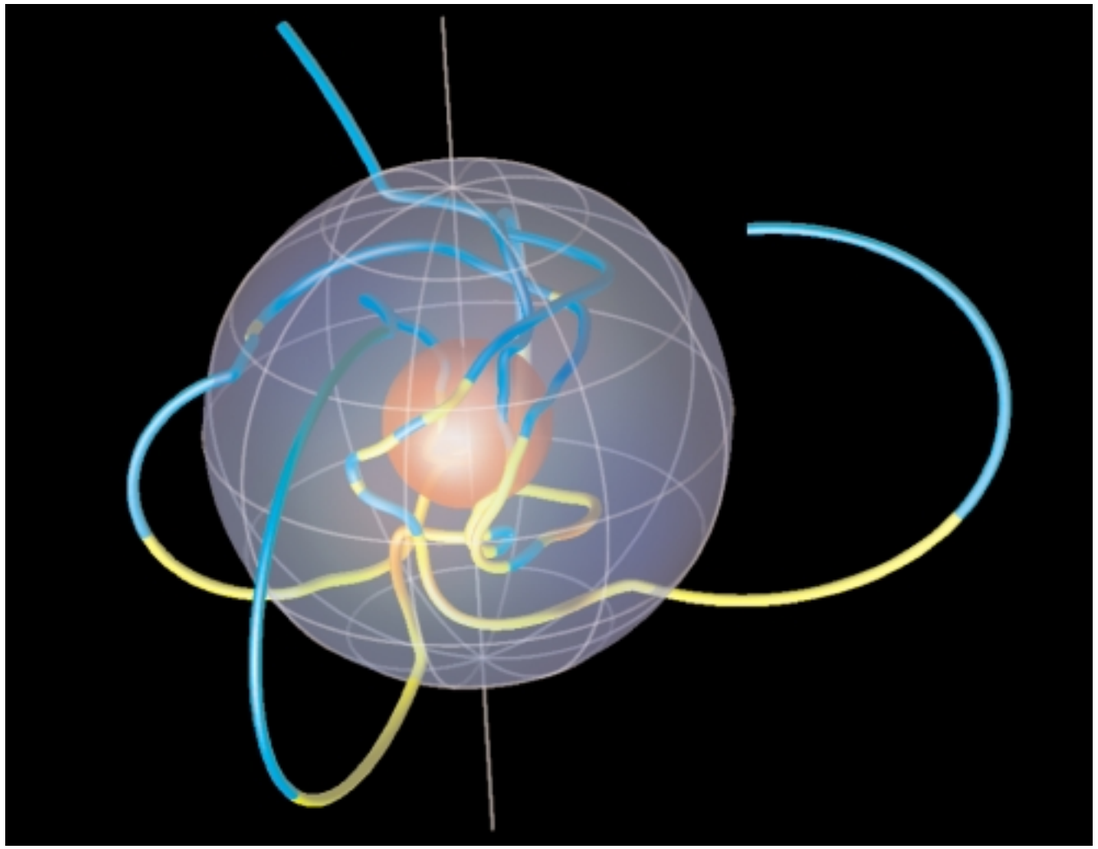
Only recently have computers been able to achieve the kinds of speeds necessary to perform numerical simulations of core dynamics. "The [dynamics of the core] are so complicated," says Weijia Kuang, "that even if you use numerical methods, you may not be able to solve the equations with the minimum required accuracy." Now, with supercomputers at centers like the NCCS, such studies are possible.

A simulation running on a typical workstation with 512 MB of RAM could produce 1,500 years of data in 1 day, whereas the Cray T3E could generate 100-fold more data in that same time.

Snapshot of one magnetic field line across Earth obtained with numerical modeling. The purple surface indicates the core–mantle boundary (CMB) that separates the fluid outer core (under) and the solid mantle (above). The red sphere in the center is Earth's solid inner core. Earth's rotation axis is indicated by the vertical white line (i.e., the top is the North Pole and the bottom is the South Pole). The blue segments of the field line show that the line is approaching the center of Earth; the orange segments show that the field line is leaving the center. The twisting and bending of the field line inside the outer core show the effect of the core flow on the magnetic field.

SOURCES OF DATA

Many secondary sources must be tapped for information: Seismology has uncovered the density, phase, temperature, and pressure profiles of the layers of the inner Earth by observing the behavior of earthquake shockwaves traveling from one spot on the surface of the globe through the interior to other places on the surface. High-pressure mineralogical studies provide information about the chemical composition of the core. Geodetic studies at the surface show the effect of the core on Earth's rotation variation. Geomagnetic observations provide information on the part of the magnetic field, the poloidal field, generated within the core in the geodynamo processes and on the variation of the field in timescales of up to hundreds of years. Radioactive dating of magnetized rocks formed at different points in Earth's history provides information about Earth's magnetic field long ago.



Early simulations showed that the physics of the model leads to the formation of magnetic North and South Poles, as humans have observed since the first compasses were invented. More importantly, the simulations can lead to magnetic field reversals on the correct timescale—an indication to Kuang and his collaborators that they are on the right track.

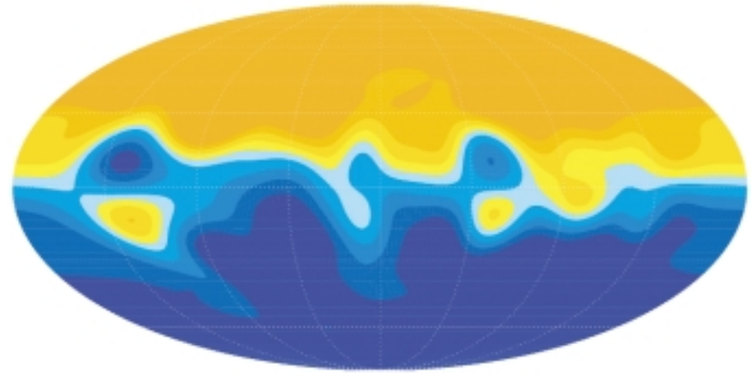
Kuang et al. have several agendas for their model, including developing a generalized version that will be made available to the public so that others might use it to probe problems related to the magnetic field. Astrophysicists, for example, might want to run the model in conjunction with studying other planets, such as Jupiter, that also have internal magnetic fields.

Geodesists such as those at GSFC use the model to study how the magnetic field affects Earth's rotation and the geopotential field on the timescales of decades. Scientists have observed slight changes in rotation for the last several decades, and believe that these changes result from the angular momentum exchange among the solid mantle and the fluid parts of the planet, including the atmosphere, the oceans, and the outer core. Additionally, geophysicists are interested in the correlation between the geodynamo and plate tectonics, and what the model can demonstrate about Earth's interior.

Kuang is interested first in understanding the details of the geodynamo dynamics—how the field is generated, how the core flow interacts with the field, and what the actions are, down to a fine level of detail. Then he can begin to address other applications of the model, such as forecasting geomagnetic field variations and understanding the dynamics of the core–mantle interaction and the properties of that boundary region, which is perhaps several hundred kilometers thick.

Of course, no model is ever perfect. In the case of the geodynamo model, many things remain to be explored, understood, and improved.

"We are still not able to simulate the real Earth because certain parameters cannot be modeled accurately due to computational limitations," says Kuang. "So far, we are still doing qualitative studies, trying to understand the phenomenon and to get the major features right."



Snapshot of the radial component of the magnetic field at the core–mantle boundary from numerical modeling. The magnetic field is shown leaving the core (orange regions) and approaching the core (blue regions).

Snapshot of the fluid density anomalies in Earth's outer core obtained from numerical modeling. The lighter (hotter) fluid parcels are shown in orange and the heavier (colder) fluid parcels in blue. The motion inside the outer core is that lighter fluid parcels rise from the inner core boundary (upwelling flow), and the heavier fluid parcels descend from the core–mantle boundary to the center of the core. The ring shapes of the fluid parcels demonstrate that there are strong zonal jets in the fluid outer core.

THE GEODYNAMO THEORY

The geodynamo theory is simple. "Basically," explains Weijia Kuang, "the Earth's outer core is liquid iron, which is electrically conducting. The motion of this liquid in a magnetic field can generate a current that, in turn, induces a magnetic field. The dynamo action is established in the core if the induced field (the magnetic field we observe at the surface of the Earth) can be sustained by the flow."

Earth's core is made of iron-rich alloy. The outer part of the core is molten, with temperatures exceeding 8,000°F, while the inner core is solid with higher iron concentration. Like any other liquid, the molten iron in the outer core flows.

Earth has been cooling slowly ever since its formation. Consequently, the solid inner core grows slowly as the heavier iron solidifies, leaving lighter constituents in the fluid outer core. The gravitational energy that is released in these processes drives the motion of the molten iron.

Moving a conductor, such as molten iron, across a magnetic field produces electrical currents, and, according to Ampere's law, electrical currents induce magnetic fields. The magnetic field can then influence the flow in the core and vice versa, in a sort of feedback loop.

The geodynamo theory posits that Earth's magnetic field is generated and maintained by the flow in Earth's outer core.

APPENDIXES

NCCS FY99 Research Projects and Principal Investigators

3–D Chemical Modeling
Ronald Prinn
Massachusetts Institute of Technology

3–D Chemistry and Transport
Richard Rood
GSFC Data Assimilation Office

3–D Chemistry and Transport Model
Anne Douglass
GSFC Atmospheric Chemistry and Dynamics Branch

4–D Data Assimilation/ TOVS Pathfinder
George Serafino
GSFC Global Change Data Center

5-Year Analysis Production
Richard Rood
GSFC Data Assimilation Office

AMIP Experiment
Richard Rood
GSFC Data Assimilation Office

AVIRIS Radiation Studies
Warren Wiscombe
GSFC Climate and Radiation Branch

Aircraft Impact Assessment
Richard Rood
GSFC Data Assimilation Office

Aircraft Winds in GLA Analysis
Joel Tenenbaum
State University of New York, Purchase

Analysis Diagnostics
Siegfried Schubert
GSFC Data Assimilation Office

Analysis Verification and Experimentation
Richard Rood
GSFC Data Assimilation Office

Analysis and Interpretation of Satellite Magnetic Data
Coerte Voorhies
GSFC Geodynamics Branch

Analysis for the Climatology and Short-Term Variability of the Atmospheric General Circulation With the GLA GEOS Data: Global Hydrological and Energy Cycles
Tsing–Chan Chen
Iowa State University

Atmospheric Remote Sensing
Yoram Kaufman
GSFC Climate and Radiation Branch

Biosphere–Atmosphere Interactions
George Collatz
GSFC Biospheric Sciences Branch

Bubbles and Drops
Ram Subramanian
Clarkson University

Chinks in the Solar Dynamo Theory
Edward DeLuca
Smithsonian Astrophysical Observatory

Climate Change
Donald Johnson
University of Wisconsin

Climate Diagnostics
Siegfried Schubert
GSFC Data Assimilation Office

Climate Modeling and Simulation
Jagadish Shukla
George Mason University and IGES/COLA

Climatic Effects of Volcanic Eruptions
Alan Robock
Rutgers University

Climatology of TOVS Water Vapor
Dennis Chesters
GSFC Mesoscale Atmospheric Processes Branch

Combined Retrieval Algorithms
James Weinman
GSFC Microwave Sensors Branch

Computation and Data Analysis
Edward Sullivan
GSFC Laboratory for Astronomy and Solar Physics

Coupled DRB Model
Watson Gregg
GSFC Oceans And Ice Branch

Crustal Dynamics
David Smith
GSFC Laboratory for Terrestrial Physics

Crustal Geomagnetic Fields
Herbert Frey
GSFC Geodynamics Branch

Cumulus Cloud Modeling
Wei–Kuo Tao
GSFC Mesoscale Atmospheric Processes Branch

DOSE and LAGEOS Investigations
Ronald Kolenkiewicz
GSFC Space Geodesy Branch

Data Analysis and Distribution
Richard Rood
GSFC Data Assimilation Office

Data Preprocessing–Preparation
Arlindo da Silva
GSFC Data Assimilation Office

Diabatic Dynamic Initialization
Michael Fox–Rabinovitz
GSFC Data Assimilation Office

Diurnal Response of Boundary
David Randall
Colorado State University

Drop Collision and Coalescence
Gretar Tryggvason
University of Michigan

Dynamical Stratospheric Model
Richard Rood
GSFC Data Assimilation Office

EGRET Data Analysis
Teresa Sheets
GSFC Data Management and Programming Office

Ecosystem Simulation Analysis
John Walsh
University of South Florida

Effect of Variations in Pacific Heat and Fresh Water Fluxes on the Indian Ocean
Roxana Wajsowicz
University of Maryland

Effects of Gravity on Sheared and Nonsheared Turbulent Nonpremixed Flames
Said Elghobashi
University of California, Irvine

Effects of Gravity on Sheared Turbulence Laden With Droplets or Bubbles
Said Elghobashi
University of California, Irvine

Energy Diagnostics
Donald Johnson
University of Wisconsin

Estimate Ocean/Air Fluxes
Shu–Hsien Chou
GSFC Mesoscale Atmospheric Processes Branch

Extra-Solar Planet Modeling
Daniel Gezari
GSFC Infrared Astrophysics Branch

Filtration for Microgravity Applications: 1) Smoldering, 2) Combustion Synthesis of Advanced Material
Bernard Matkowsky
Northwestern University

Flame Spread in Non-Uniform Mixtures
Fletcher Miller
NASA, John H. Glenn Research

Center

Flame Spread on Liquid Surface
William Sirignano
University of California, Irvine

Fractal Cloud Structure
Robert Cahalan
GSFC Climate and Radiation Branch

GEODYN Software Development
David Rowlands
GSFC Space Geodesy Branch

GLA TOVS Satellite Data—Development, Processing, Validation, and Scientific Research
Joel Susskind
GSFC Laboratory for Atmospheres

GMAS Hydrology Parameterization
Michael Jasinski
GSFC Hydrological Sciences Branch

GPS Analysis
Braulio Sanchez
GSFC Space Geodesy Branch

GSFC Cumulus Cloud Modeling
Joanne Simpson
GSFC Mesoscale Atmospheric Processes Branch

GSFC Earth Science Data and Information Systems (ESDIS) Program
Christopher Bock
GSFC Science Data Systems Branch

Gas-Phase Combustion Synthesis of Metal and Ceramic Nano-Particles
Richard Axelbaum
Washington University of St. Louis Missouri

General Analysis
Franklin Ottens
GSFC Laboratory for Extraterrestrial Physics

Geomagnetic Field Studies
Herbert Frey
GSFC Geodynamics Branch

Geopotential Determination Using Satellite, Gravity, and Ocean Model Data
Byron Tapley
University of Texas at Austin

Global Geodynamics
Benjamin Chao
GSFC Space Geodesy Branch

Global Inventory Mapping and Monitoring (GIMMS)
Compton Tucker
GSFC Biospheric Sciences Branch

Global Model Studies
George Emmitt
Simpson Weather Associates

Global Modeling of Tropospheric Chemistry
Daniel Jacob
Harvard University

Global Tides from Altimetry
Braulio Sanchez
GSFC Space Geodesy Branch

Gravity Model Development
Frank Lemoine
GSFC Space Geodesy Branch

HEASARC
Teresa Sheets
GSFC Data Management and Programming Office

HST Phase Retrieval
Howard Wood
GSFC HST Flight Systems & Servicing Project

High Performance Computing
Richard Rood
GSFC Data Assimilation Office

IAU and Model/Analysis Interface
Richard Rood
GSFC Data Assimilation Office

IMP-J Magnetic Field Experiment
Ronald Lepping
GSFC Electrodynamics Branch

ISEE-1 VES Data Analysis
Keith Ogilvie
GSFC Laboratory for Extraterrestrial Physics

ISEE-OGH Data Analysis
Keith Ogilvie
GSFC Laboratory for Extraterrestrial Physics

Imaged EUV Spectra from SERTS
Roger Thomas
GSFC Solar Physics Branch

Improving Land Hydrologic Processes
Yogesh Sud
GSFC Climate and Radiation Branch

Improving Radiation Codes
Albert Arking
Johns Hopkins University

Inference of Global Warming from Satellite Data
Prabhakara Cuddapah
GSFC Climate and Radiation Branch

Influence of Land Surface Processes/Land Cover Changes in Amazon Regional Hydrometeorology
Yongkang Xue
University of Maryland and Univeristy of California, Los Angeles

Instability of Alfven Wave
Adolfo Viñas
GSFC Interplanetary Physics Branch

Integrating Biogeochemical, Ecological, and Hydrological Processes in a Dynamic Biosphere
Jonathan Foley
University of Wisconsin

Interannual Variability in Biochemical Cycles of the North Atlantic
John Marshall
Massachusetts Institute of Technology

Interdisciplinary Science Investigations
David Randall
Colorado State University

Investigation of GRXE Production
Teresa Sheets
GSFC Data Management and Programming Office

Kalman Filtering
Richard Rood
GSFC Data Assimilation Office

Lidar
James Spinhirne
GSFC Mesoscale Atmospheric Processes Branch

Land Climatology
Yogesh Sud
GSFC Climate and Radiation Branch

Land Surface Component of the Climate System
Randal Koster
GSFC Hydrological Sciences Branch

Liquid He Heat Studies
Efstratios Manousakis
Florida State University

Low-Frequency Phenomenon Modeling
William Lau
GSFC Climate and Radiation Branch

MAGSAT Crustal Anomalies
Patrick Taylor
GSFC Geodynamics Branch

MHD Turbulence
Melvyn Goldstein
GSFC Interplanetary Physics Branch

Magnetic Wave–Particle Interaction
Richard Denton
Dartmouth College

Magnetospherical Simulation
Steven Curtis
GSFC Planetary Magnetospheres Branch

Mars Observer Gamma Ray Spectrometer
Jacob Trombka
GSFC Astrochemistry Branch

Mars Observer Laser Altimeter
David Smith
GSFC Laboratory for Terrestrial Physics

Mesoscale Dynamics
Wei-Kuo Tao
GSFC Mesoscale Atmospheric Processes Branch

Mid-Latitude Ocean Circulation
David Adamec
GSFC Oceans And Ice Branch

Miscellaneous Analysis Production
Richard Rood
GSFC Data Assimilation Office

Model Diagnostics
Siegfried Schubert
GSFC Data Assimilation Office

Model Production
Richard Rood
GSFC Data Assimilation Office

Model Research
Mark Helfand
GSFC Data Assimilation Office

Model and STRATAN Analysis
Richard Rood
GSFC Data Assimilation Office

Modeling Studies of LFV
Randall Dole
National Oceanic and Atmospheric Administration

Modeling of Decadal Variability in the Tropical Atlantic Climate
Vikram Mehta
GSFC Climate and Radiation Branch

Molecular Dynamics
Joel Koplik
City College of the City University of New York

Momentum and Energy Budgets in Models
David Salstein
AER, Inc.

NAOS Studies
Stephen Lord
University of Maryland

Nimbus SMMR Data
Milton Halem
GSFC Earth and Space Data Computing Division

Non-local Coupling by Magnetospheric Currents
Keith Siebert
Mission Research Corporation

North Pacific Circulation Modeling
Michele Rienecker
GSFC Oceans And Ice Branch

Numerical Simulations in Support of SOHO
Arthur Poland
GSFC Solar Physics Branch

Numerical Simulations/Cloud Studies
David Randall
Colorado State University

Objective Analysis Methods
Grace Wahba
University of Wisconsin

Ocean Surface Wind Analysis
Robert Atlas
GSFC Data Assimilation Office

Oceans and Ice Branch General Access
Michele Rienecker
GSFC Oceans And Ice Branch

Pioneer/Helios Data Analysis
Teresa Sheets
GSFC Data Management and Programming Office

Planetary Scale Interactions
William Lau
GSFC Climate and Radiation Branch

Polar Ocean Modeling Studies
Sirpa Häkkinen
GSFC Oceans And Ice Branch

Radiation Budget
Man-Li Wu
GSFC Data Assimilation Office

Radiation Budgets/Aerosol Radiative Forcing
Ming-Dah Chou
GSFC Climate and Radiation Branch

Rain Retrieval Studies
Prabhakara Cuddapah
GSFC Climate and Radiation Branch

Rainfall Assimilation Studies
Vijaya Karyampudi
GSFC Mesoscale Atmospheric Processes Branch

Reduction & Analysis of LEO Satellite Magnetic Field Data
Joseph Cain
Florida State University

Regional Application Centers
Patrick Coronado
GSFC Applied Information Sciences Branch

Remote Sensing of Clouds
Si-Chee Tsay
GSFC Climate and Radiation Branch

Research/Student Mentor Program
Jan Hollis
GSFC Earth and Space Data Computing Division

Resonances
Anand Bhatia
GSFC Solar Physics Branch

STIS GTO Research Support Using UniTree Archival Data
Sara Heap
GSFC UV Optical Astronomy Branch

Satellite Data Evaluation
Richard Rood
GSFC Data Assimilation Office

Satellite Rainfall Estimates
Man-Li Wu
GSFC Data Assimilation Office

Science Network Office
J. Patrick Gary
GSFC Earth and Space Data Computing Division

Search and Rescue SAR Research
David Affens
GSFC Microwave Systems Branch

Seasonal-to-Interannual Prediction
Michele Rienecker
GSFC Oceans and Ice Branch

Seasonal-to-Interannual Climate Variability
Michael Ghil
University of California, Los Angeles

Seasonal-to-Interannual Collaboration
Michele Rienecker
GSFC Oceans and Ice Branch

Semi-LaGrangian GCM Development
Yong Li
GSFC Data Assimilation Office

Signatures at Magnetopause Reconnection
George Siscoe
Boston University

Simulation Studies
Robert Atlas
GSFC Data Assimilation Office

Simulation of the Global Radiation Impact of Anthropogenic Aerosols
Steven Ghan
Batelle Pacific Northwest National Laboratory

Solar Activity
Gordon Holman
GSFC Solar Physics Branch

Solar System and Milankovitch Cycles
David Rubincam
GSFC Geodynamics Branch

Space and Solar Plasmas Simulations
Ken-Ichi Nishikawa
Louisiana State University

Stratosphere Analysis Production
Richard Rood
GSFC Data Assimilation Office

Stratospheric Analysis Research
Richard Rood
GSFC Data Assimilation Office

Structural Chemistry
Jerome Karle
Naval Research Laboratory

TOGA/COARE Atmospheric Circulation
Winston Chao
GSFC Climate and Radiation Branch

TOMS and SBUV-2 Data Processing
Jay Herman
GSFC Atmospheric Chemistry and Dynamics Branch

TOPEX Software Development and Operations
Benjamin Chao
GSFC Space Geodesy Branch

TRMM Rain Climatology, Errors and Models
Thomas Bell
GSFC Climate and Radiation Branch

Time Variable Gravity
Frank Lemoine
GSFC Space Geodesy Branch

Tropical Climate Variability
William Lau
GSFC Climate and Radiation Branch

Tropical Cyclone Rainfall
Edward Rodgers
GSFC Mesoscale Atmospheric Processes Branch

Tropical Ocean Circulation
Antonio Busalacchi
GSFC Laboratory for Hydrospheric Processes

Tropical Rainfall and Climate
William Lau
GSFC Climate and Radiation Branch

Tropospheric Chemistry Aerosol Model
Mian Chin
GSFC Atmospheric Chemistry and Dynamics Branch

Tropospheric Convection and Stratosphere-Troposphere Exchange
Kenneth Pickering
GSFC Atmospheric Chemistry and Dynamics Branch

Tropospheric Photochemistry
Anne Thompson
GSFC Atmospheric Chemistry and Dynamics Branch

Type III Solar Radio Bursts in the Corona
Iver Cairns
Naval Research Laboratory

UIT Data Reduction Pipeline
Susan Neff
GSFC UV Optical Astronomy Branch

Using NASA Full Physics Adjoint of the GEOS–1 GCM for Retrospective Analysis
Ionel Navon
Florida State University

Version 1 Research and Development
Arlindo da Silva
GSFC Data Assimilation Office

Version 2 Research and Development (Global OI)
Arlindo da Silva
GSFC Data Assimilation Office

Voyager Magnetometer Data
Mario Acuña
GSFC Planetary Magnetospheres Branch

Voyager Neptune Atmosphere Composition
William Maguire
GSFC Planetary Systems Branch

Voyager Plasma Data Analysis
Keith Ogilvie
GSFC Laboratory for Extraterrestrial Physics

Water Vapor and Cloud Feedback
Albert Arking
Johns Hopkins University

X-Ray Data Analysis
Teresa Sheets
GSFC Data Management and Programming Office

X-Ray Study of 1E1740
Joseph Dolan
GSFC UV Optical Astronomy Branch

Acronyms, Names, and Abbreviations

ACARS	Aircraft Communications, Addressing, and Reporting System
AFRCC	Air Force Rescue Coordination Center
AGCM	atmospheric general circulation model
AMIP	Atmospheric Model Intercomparison Project
AMS	American Meteorological Society
ARTCC	Air Route Traffic Control Center
atexpert	autotask expert
ATM	asynchronous transfer mode
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
BLAS	basic linear algebra subroutines
C	a high-level programming language
C++	a high-level object-oriented programming language
ccNUMA	cache-coherent Non-Uniform Memory Access
CFAR	constant false alarm rate
CHAMP	a German satellite mission for gravity field mapping
CLARiiON	a type of storage disk manufactured by EMC Corp.
CMB	core–mantle boundary
CMC	Canadian Meteorological Center
COARE	Coupled Ocean–Atmosphere Response Experiment
COLA	Center for Ocean–Land–Atmosphere Studies
COSPAS	the Russian-operated equivalent of SARSAT
CPU	central processing unit
CU	computing unit
CUC	Computer Users Committee
CUNY	City College of New York
DAO	Data Assimilation Office
DOSE	Dynamics of the Solid Earth
ECMWF	European Center for Medium Range Weather Forecasts
EGM96	Earth Gravitational Model 1996
EGRET	Energetic Gamma-Ray Experiment Telescope
EISPACK	eigenvalues analysis package
ELT	emergency locator transmitters
Emacs	editor macros
EMC	a type of storage disk manufactured by EMC Corp.
EnKF	ensemble Kalman filter
ENSO	El Niño Southern Oscillation
EOS	Earth Observing System
ESA	European Space Agency
ESDCD	Earth and Space Data Computing Division
ESDIS	Earth Science Data and Information Systems program
EUV	Extreme Ultraviolet spectrometer
F	Fahrenheit
FAA	Federal Aviation Administration
FDDI	fiber distributed data interface
FISHPACK	Fortran subprograms for separable elliptic partial differential equations
FLINT	Fortran LINT
Fortran	formula translation
FY	fiscal year
GB	gigabyte/billion bytes
GB/sec	billion bytes per second
GCM	general circulation model
GEODYN	Geodynamics Orbit And Geodetic Parameter Estimation System
GEOS	Goddard Earth Observing System
GEOS–1	the first Goddard Earth Observing System model
GFLOPS	billion floating-point operations per second
GFZ	GeoForschungsZentrum
GIMMS	Global Inventory Mapping and Monitoring System
GIS	Geographic Information System
GLA	Goddard Laboratory for Atmospheres
GMAS	Goddard Mesoscale Atmospheric Simulation system
GMT	Greenwich mean time
GNU	GNUs Not Unix
GOCE	Gravity Ocean Climate Experiment
GOES	Geostationary Operational Environmental Satellite
GPLLJ	Great Plains low-level jet
GPS	Global Positioning System
GRACE	Gravity Recovery And Climate Experiment

GRXE	Galactic ridge x-ray emissions
GSFC	Goddard Space Flight Center
GSM	Global Spectral Model
hdf	hierarchical data format
HEASARC	High Energy Astrophysics Science Archive Research Center
HIRS	High-Resolution Infrared Sounder
HP	Hewlett-Packard
hpm	hardware performance monitor
HPCC	High-Performance Computing and Communications
HST	Hubble Space Telescope
HiPPI	High-Performance Parallel Interface
I/O	input/output
IBM	International Business Machines Corp.
IGES	Institute of Global Environment and Society, Inc.
IMF 4	an interseasonal oscillation index
IMP	Interplanetary Monitoring Platforms
IMSL	International Mathematical and Statistical Libraries
IP	Internet protocol
ISEE	International Sun–Earth Explorer program
IT	information technology
JPL	Jet Propulsion Laboratory
LAGEOS	laser geodynamics satellite
LASP	Laboratory for Astronomy and Solar Physics
LEO	Low Earth Orbiting
LEP	Laboratory for Extraterrestrial Physics
LFV	Lower Fraser River Valley
lidar	laser radar
LINT	a degugging tool that checks C and C++ source code for non-portability
LLJ's	low-level jets
MAGSAT	Magnetic Field Satellite
MAX	Mid-Atlantic Crossroads
MB	megabyte/million bytes
MB/sec	million bytes per second
MCC	Mission Control Center
MD	molecular dynamics
MHD	magnetohydrodynamics
MHz	megahertz
MJO	Madden and Julian Oscillation
MSFC	Marshall Space Flight Center
MW	megaword
Mbps	million bits per second
NAG	Numerical Algorithms Group
NAOS	North American Observing Systems Program
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCCS	NASA Center for Computational Sciences
NCEP	National Centers for Environmental Prediction
NIMA	National Imagery and Mapping Agency
Nimbus	A series of seven satellites launched by NASA between 1958 and 1978
NOAA	National Oceanic and Atmospheric Administration
NREN	NASA Research and Engineering Network
NSCAT	NASA scatterometer
NSIPP	NASA's Seasonal-to-Interannual Prediction Project
NUMA	Non-Uniform Memory Access
NWS	National Weather Service
OC	optical carrier
OC–12	622.08 Mbps
ODEPACK	ordinary differential equations package
OGCM	ocean general circulation model
OLR	Outgoing Longwave Radiation
PB	petabyte/10^{15} bytes
perftrace	performance trace
prof	profile
QuikSCAT	NASA Quick Launch Scatterometer
qv	moisture transport
R&D	research and development
RCC	Rescue Coordination Center
RUC	Rapid Update Cycle
Raytheon ITSS	Raytheon Information Technology and Scientific Services
S&R	search and rescue
SAN	Science ATM Network
SAR	synthetic aperture radar
SAR^2	synthetic aperture radar search and rescue
SARSAT	Search and Rescue Satellite-Aided Tracking
SBUV	Solar Backscatter Ultraviolet
SCB	Science Computing Branch
SCSI	small computer system interface
SERTS	Solar EUV Research Telescope and Spectrograph
SLATEC	Sandia, Los Alamos Air Force Weapons Laboratory Technical Exchange Committee
SMMR	Scanning Multichannel Microwave Radiometer
SOHO	Solar and Heliospheric Observatory
SSH	sea surface height
SST	sea surface temperature
SSU	Stratospheric Sounding Unit
STK	StorageTek
SUNY	State University of New York
T&EWG	Test and Evaluation Working Group
TAG	Technical Assistance Group
TAO	the Tropical Atmosphere Ocean project
TB	terabyte/10^{12} bytes
TCP	Transmission Control Protocol
TDRSS	Tracking and Data Relay Satellite System
TFLOPS	trillion floating-point operations per second
TIROS	Television-Infrared Observation Satellite
TOGA	Tropical Ocean Global Atmosphere
TOMS	Total Ozone Mapping Spectrometer
TOPEX/Poseidon	US–French ocean topography experiment satellite
totalview	a parallel debugger
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
UAV	Unmanned Aerial Vehicle
UCLA	University of California, Los Angeles
UCSC	University of California, Santa Cruz
UIT	Ultraviolet Imaging Telescope
UNICOS	Unix Cray Operating system
UNIX	an operating system based on the C programming language
UTSI	UniTree Software, Inc.
UV	ultraviolet
UniTree	a hierarchical data archival system
VCA	Variance Component Analysis
VES	viscoelastic surfactant
vi	visual editor
WWW	World Wide Web

Researchers' Affiliations

David W. Affens 76
GSFC Microwave Systems Branch

Robert M. Atlas 28, 37
GSFC Data Assimilation Office

Jayanth Banavar 64
The Pennsylvania State University

Benjamin Fong Chao 90
GSFC Space Geodesy Branch

Roy Dreibelbis 76
Computer Sciences Corporation

Mark Helfand 31
GSFC Data Assimilation Office

Steven Klosko 83
Raytheon ITSS

Joel Koplik 64
City College of the City University of New York

Weija Kuang 90
GSFC Space Geodesy Branch

Frank Lemoine 84
GSFC Space Geodesy Branch

Stephen Lord 36
National Oceanic and Atmospheric Administration

Arthur Mansfield 76
GSFC Microwave Systems Branch

Erricos Pavlis 84
University of Maryland, Baltimore County, and the GSFC Space Geodesy Branch

Nikos Pavlis 84
Raytheon ITSS

Michele Rienecker 20
GSFC Oceans And Ice Branch

Braulio Sanchez 69
GSFC Space Geodesy Branch

Thomas Schlatter 35
National Oceanic and Atmospheric Administration

Siegfried Schubert 29
GSFC Data Assimilation Office

Jagadish Shukla 41
George Mason University and the Center for Ocean–Land–Atmosphere Studies

Joanne Simpson 53
GSFC Mesoscale Atmospheric Processes Branch

Max Suarez 20
GSFC Climate and Radiation Branch

Joel Susskind 49
GSFC Sounder Research Team

Man–Li Wu 33
GSFC Data Assimilation Office

Acknowledgements

Production

Jason Bardi
Carol Boquist
Ronnie Estes
Carol Ladd
John Hazen
Judy Laue

Management

Michael Seablom
Susan Kaltenbaugh

In addition to the individuals interviewed for these articles, many others contributed their time and effort:

Jeanne Beatty
Timothy Burch
Nicholas Burke
Lara Clemence
Thomas Clune
Jarrett Cohen
Harold Domchick
Igor Eberstein
Richard Glassbrook
Katharine Gurski
Susan Hart
Victoria Jenkins
Medora Macie
Michael McGunigale
Peter Norris
Nancy Palm
Sherri Panciera
Michael Purucker
George Rumney
Marvin Saletta
Thomas Schardt
Jeff Staddon
Spencer Swift
Adina Tarshish
James Taylor

Additional Photo Credits

Cover, 3–6, Jason Bardi
Cover, iv, v, 77, 94–95, NASA

If you would like more information about the NCCS or would like to obtain a copy of this document, please contact:
tag@nccs.gsfc.nasa.gov

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland

NP-2000-8-490-GSFC